



Cite this: DOI: 10.1039/d0cs01551b

# Halogenases: a palette of emerging opportunities for synthetic biology–synthetic chemistry and C–H functionalisation

Charlotte Crowe,<sup>†‡</sup> Samuel Molyneux,<sup>‡</sup> Sunil V. Sharma, Ying Zhang, Danai S. Gkotsi,<sup>§</sup> Helen Connaris and Rebecca J. M. Goss<sup>id\*</sup>

The enzymatic generation of carbon–halogen bonds is a powerful strategy used by both nature and synthetic chemists to tune the bioactivity, bioavailability and reactivity of compounds, opening up the opportunity for selective C–H functionalisation. Genes encoding halogenase enzymes have recently been shown to transcend all kingdoms of life. These enzymes install halogen atoms into aromatic and less activated aliphatic substrates, achieving selectivities that are often challenging to accomplish using synthetic methodologies. Significant advances in both halogenase discovery and engineering have provided a toolbox of enzymes, enabling the ready use of these catalysts in biotransformations, synthetic biology, and in combination with chemical catalysis to enable late stage C–H functionalisation. With a focus on substrate scope, this review outlines the mechanisms employed by the major classes of halogenases, while in parallel, it highlights key advances in the utilisation of the combination of enzymatic halogenation and chemical catalysis for C–H activation and diversification.

Received 14th December 2020

DOI: 10.1039/d0cs01551b

rsc.li/chem-soc-rev

School of Chemistry, and BSRC, University of St Andrews, North Haugh, St Andrews KY16 9ST, UK. E-mail: rjmg@st-andrews.ac.uk

<sup>†</sup> Now studying at: Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, UK.

<sup>‡</sup> These authors contributed equally.

<sup>§</sup> Now carrying out postdoctoral research at: School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, West Yorkshire, LS2 9JT, UK.

## Introduction

Halogenated organic compounds play an important role in society. In the agrochemical sector, a staggering 96% of herbicides, fungicides, insecticides, acaricides and nematocides produced since 2010 contain halogen atoms.<sup>1</sup> Such compounds are valuable commodities; for example, the top-selling insecticide imidacloprid (Fig. 1A) **1** had sales of 1.09 billion USD in 2011 alone,<sup>2</sup> and in 2013, 50% of marketed top drugs contain



Charlotte Crowe

Charlotte Crowe received her MChem in Chemistry (2020) from the University of St Andrews. She undertook a literature project in the area of halogenases and a laboratory project in organic synthesis performing mild-aqueous cross-coupling, both under the supervision of Rebecca Goss. She is currently a PhD student in the laboratory of Alessio Ciulli at the University of Dundee. Her interdisciplinary research focuses on the mechanisms of targeted protein degradation mediated by small molecule degraders such as PROTACs and molecular glues.



Samuel Molyneux

Samuel Molyneux is a PhD student under the supervision of Professor Rebecca Goss at the University of St Andrews and is part of and funded by the Critical Resource Catalysis (CRITICAT) CDT. He earned his Master's and Bachelor's degree from Trinity Hall, Cambridge in 2018, and his current research interests within the Goss group include the merger of homogeneous metal catalysis with biocatalysis in aqueous systems.



halogen atoms.<sup>3,4</sup> The fluorinated hepatitis C ledipasvir/sofosbuvir combination (Harvoni<sup>®</sup>) **6** was the second best-selling blockbuster pharmaceutical in 2016 with USA sales revenues of 4.9 billion USD (Fig. 1B),<sup>5</sup> whilst, in the same year, US sales of the chlorinated antihistamine loratadine (Claritin<sup>®</sup>) **11** exceeded 2 billion USD.<sup>6</sup> An economic report analysing the annual sales of the 100 top selling pharmaceuticals reveals that 12.9% of the final active pharmaceutical ingredient (API) is chlorinated or brominated, whilst a further 62.7% require chlorination at some stage in their manufacture.<sup>7</sup>

The incorporation of a carbon–halogen bond is a regularly employed strategy in medicinal and agro-chemistry. Due to bond polarisation and the ability to form highly directional halogen bonds to electron-rich oxygen, nitrogen, sulfur and aromatic rings, halogen incorporation may be used to enhance stability, and increase lipophilicity which in turn can influence bioavailability and modulate compound bioactivity.<sup>8</sup> Beyond these applications, series of fluorinated and iodinated compounds are used in radiotherapy and biomedical imaging.<sup>9,10</sup> Halogenation is also central to molecule construction and manufacture. The reactivity and chemical orthogonality of organohalides enables



**Sunil V. Sharma**

*Sunil V. Sharma obtained a post-graduate degree in pharmaceutical sciences and medicinal chemistry from India and a PhD in organic chemistry from University of East Anglia, Norwich, UK. He then held post-doctoral position with Dr C. Hamilton (BBSRC/Leverhulme trust), UEA, Norwich. Sunil joined the research group of Prof. Rebecca Goss, FRSC, FRSE in St Andrews as a senior research fellow. He has expertise*

*in medicinal chemistry, synthetic methodology development and natural product research. Sunil is a recipient of the prestigious RSE Enterprise fellowship (2020) and has a keen interest in translation of (bio)-catalytic chemical processes for applications within the agrochemical and pharmaceutical industries.*



**Ying Zhang**

*Ying Zhang is a PhD student under the supervision of Professor Rebecca Goss at the University of St Andrews and is funded by the China Scholarship Council (CSC). She received her Bachelor's and Master's degrees at Tianjin University, China. Ying has expertise in bioinformatics, heterologous expression and production of microbial proteins. Her current interests within the Goss group is in the field of biocatalysis and identifying novel halogenases as tools for synthesis of halogenated compounds.*



**Danai S. Gkotsi**

*Danai S. Gkotsi received her BSc with honours in Molecular Biology, at the University of Aberdeen, where she focused on the comparative functional genomics of ion channelopathies. In 2013, she joined the group of Professor Rebecca Goss in St Andrews, with an iCASE studentship, funded by Syngenta. This was to complete her PhD on the identification and development of halogenases as tools for synthesis. She is currently working*

*as a mechanistic cell biology fellow with Dr Riobo-Del Galdo at the University of Leeds, aiming to understand the non-canonical function of key developmental membrane receptors, as cholesterol transporters in autophagy.*

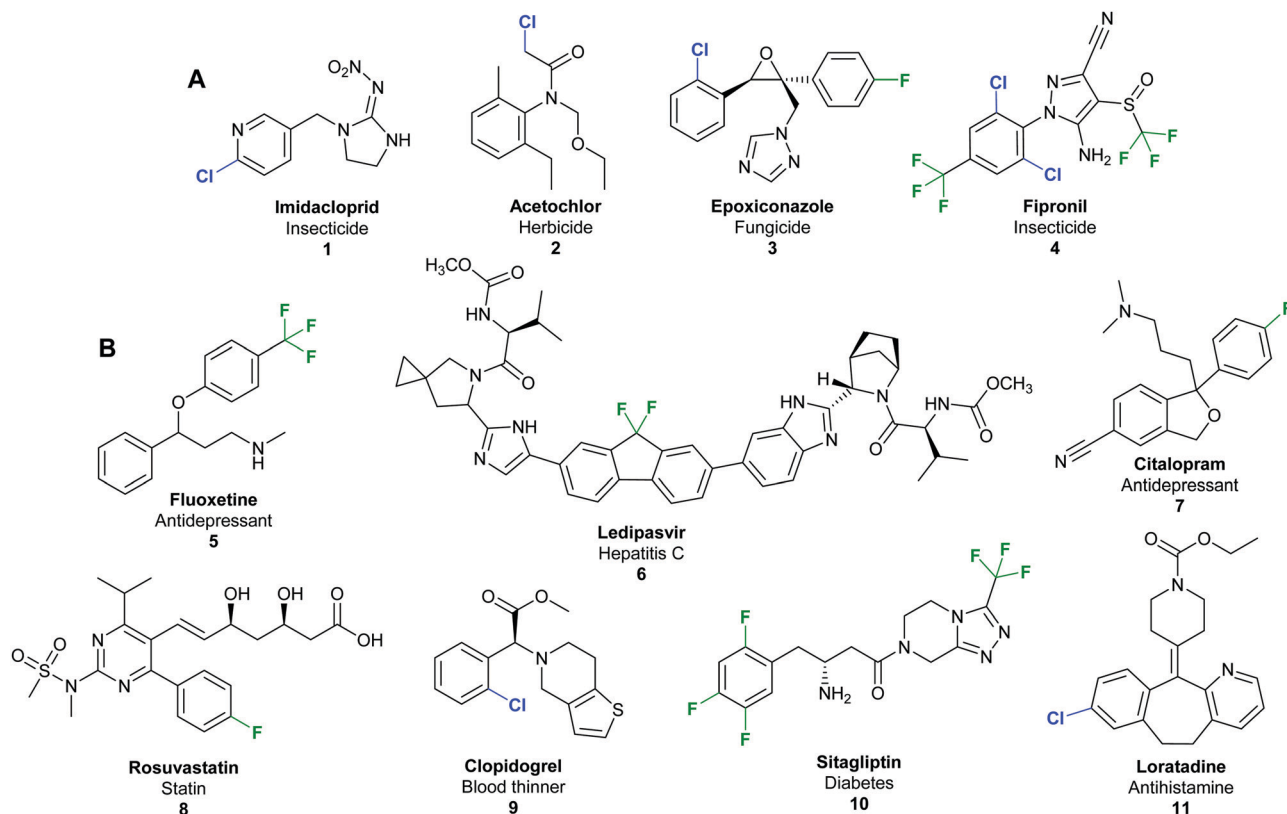


**Helen Connaris**

*Helen Connaris received her PhD in Biochemistry from University College London following a BSc (Hons) Industrial Biology degree from Southbank University. From 1992-1998, she held post-doctoral positions in London and Bath, in the field of extremophiles. She joined Professor Garry Taylor's group at St Andrews in 1999, where she specialized in developing novel proteins using microbial carbohydrate-binding domains (CBMs) and*

*was the first to recognize their therapeutic potential for disease. In 2016, Helen became Principal Investigator, leading the CBM research team for Pneumagen, also as a Founder and CSO. In 2020, Helen joined the group of Professor Rebecca Goss as a Senior Research Fellow, exploring novel halogenases as biocatalytic tools for halogenation of industrial-relevant compounds.*





**Fig. 1** Examples of commercially important and structurally diverse organohalides. (A) Examples of top-selling halogenated agrochemicals include imidacloprid **1**, biggest selling insecticide 2011 sales of 1.09 billion USD per annum;<sup>2</sup> acetochlor **2**, herbicide in the top 10 best-selling agrochemicals in 2003;<sup>16</sup> epoxiconazole **3**, a triazole based fungicide with sales of 215 million USD in 2001;<sup>17</sup> and fipronil **4**, a pyrazole based insecticide with sales of 270 million USD in 2001.<sup>17</sup> (B) Examples of top-selling halogenated pharmaceuticals include antidepressants fluoxetine **5** (Prozac<sup>®</sup>)<sup>17</sup> and citalopram **7** (Celexa<sup>®</sup>);<sup>18</sup> anti-hepatitis C agent ledipasvir **6**;<sup>5</sup> statin rosuvastatin **8**;<sup>19</sup> the blood thinner clopidogrel **9**;<sup>16</sup> the type 2 diabetes treatment sitagliptin **10**;<sup>20</sup> and antihistamine loratadine **11**.<sup>6</sup>

many selective transformations including, most notably, cross-coupling and substitution chemistries. Synthetically, the generation of organohalides is currently achieved through the

addition of molecular halogens, hydrohalogenation, or the use of reagents enabling nucleophilic substitution and electrophilic aromatic substitution [EAS]. These reactions are generally accompanied by environmental and selectivity challenges due to the toxic, corrosive and reactive nature of the reagents used, and the propensity to generate complex product mixtures.<sup>11</sup> Whilst recent advances in homogenous catalyst design has meant more selective catalytic approaches towards halogenation are known, such catalysts can be complex, costly and time consuming to synthesise.<sup>12</sup>

Though naturally occurring organohalides were initially speculated to be nothing more than a biological curiosity, artefact, or to be anthropogenic in origin, today, over 5000 halogenated natural products are known, with genes encoding halogenases identified across all kingdoms of life.<sup>13</sup> These represent a resource for greener and often highly regioselective halogenation. An increased appreciation of mechanisms underpinning enzymatic halogenation has evolved over the past 25 years, and may broadly be classified as electrophilic, nucleophilic or radical halogenation. Whilst a good level of understanding of the enzymatic active sites and structural requirements of these enzymes exists, a number of fundamental details remain to be elucidated.



**Rebecca Goss**

*Rebecca Goss is Professor of Organic/Biomolecular Chemistry at the University of St Andrews. Her team's research is on medically relevant bioactive compounds, particularly natural products, and in hyphenating Synthetic Biology – Synthetic Chemistry to enable access analogues of these compounds. The team also focus on discovering and developing enzymes as tools for synthesis and working toward sustainable antibiotic production*

*for Africa. She is founder of St Andrews Multidisciplinary Anti-infective Research and Therapeutics Centre (SMART Centre) and Scottish Health and Innovation Network for Entrepreneurship (SHINE).*





This review overviews the mechanisms of the different classes of halogenases, the natural context of the halogenases, their biocatalytic applications, and explores their powerful potential for use in the hyphenation of Synthetic Biology and Synthetic Chemistry (SynBio-SynChem) enabling C-H activation and diversification.

## 1. Mechanisms for biocatalytic halogenation

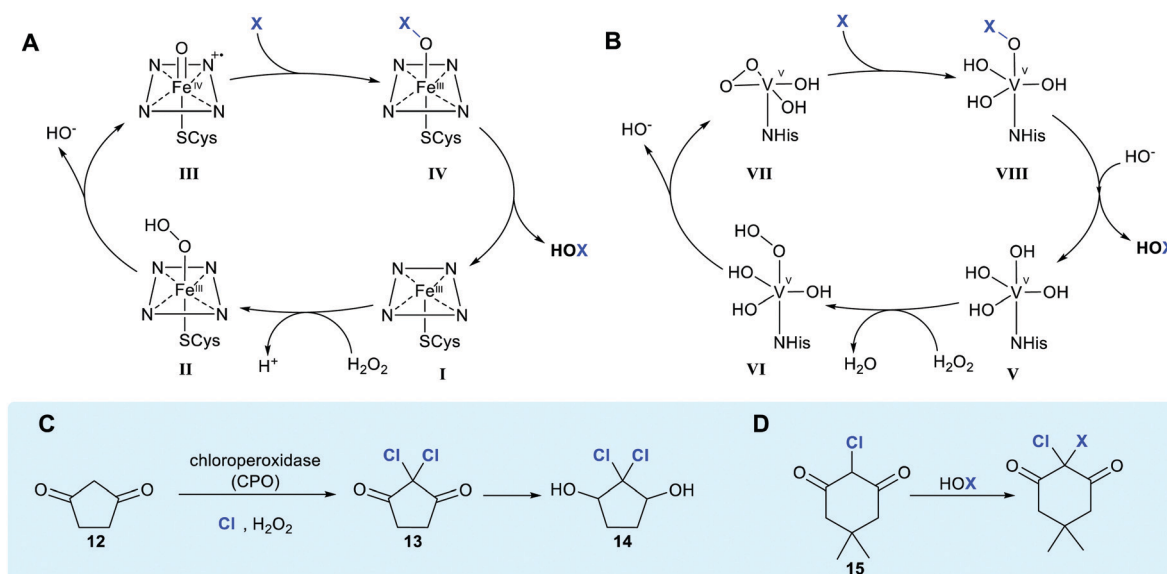
Enzymatic halogenation is classified mechanistically in accordance with whether the halogen is introduced by an electrophilic, nucleophilic, or radical halide species, with the vast majority of the enzymes that have been discovered so far, being electrophilic halogenases, mediating oxidation of the halogen to generate a reactive electrophilic species. There is further subdivision within the broader halogenase categories, based on the exact catalytic species and cofactor which enables halogen oxidation.

### 1.1 Electrophilic ( $X^+$ ) halogenation

**1.1.1 Haloperoxidases.** Haloperoxidases are defined as enzymes that utilise hydrogen peroxide and chloride, bromide or iodide to generate HOX. These enzymes are named in accordance with the most electronegative halogen for which they can affect oxidation. Haloperoxidases are generally the least regioselective of the enzymes enabling halogenation; catalysing the generation of free hypohalous acid, which in all but a few cases, is released by the enzyme. The earliest reported studies of halogenation enabled by an enzyme date back to 1959, when Lowell Hager discovered a haloperoxidase (HPO) from the fungus *Calariomyces fumago*.<sup>14</sup> The enzyme that Hager had utilised as part of a lysate was named chloroperoxidase or (CPO) and was shown to be capable of affecting dichlorination of 1,3-cyclopentanedione **12**. It has thus been postulated to be

involved in the biosynthesis of the antibiotic caldariomycin **14** (Scheme 1C).<sup>15</sup> CPO may be described more specifically as a heme iron-dependent haloperoxidase, a subset of haloperoxidases that contain a heme-iron cofactor coordinated to an axial cysteine ligand within the enzyme. Hydrogen peroxide binds axially to heme-Fe(III) in its resting state to form a  $^{\bullet+}$ heme-Fe(IV)=O intermediate, denoted compound I.<sup>23</sup> Addition of halide to the ferryl oxygen produces a heme-Fe(III)-OX species (Scheme 1A, IV). Free hypohalous acid, HOX, is generated which reacts with an electron-rich substrate (Scheme 1A).<sup>24</sup> Generation of electrophilic  $X^+$  allows halogenation of electron-rich compounds by electrophilic aromatic substitution, generally at the most reactive site(s) within the substrate. CPO has been used biocatalytically to mediate the chlorination, bromination or iodination of a variety of electron-rich substrates (see Section 2), however, as it lacks a site to position the substrate, no regiocontrol beyond that conferred by the electronics of the substrate is observed.

The vanadium-dependent haloperoxidases (V-HPOs) operate by an analogous mechanism, again generating and, mostly releasing, free hypohalous acid. This subcategory of haloperoxidase is widely found in algae, fungi and bacteria, and utilises a vanadium cofactor bound by an axial histidine ligand.<sup>25</sup> In the resting state, the cofactor is in a trigonal bipyramidal configuration (Scheme 1B, V), a distal  $H_2O_2$  displaces two hydroxyl ligands giving rise to formation of the known  $\eta^2$ -peroxido complex (Scheme 1B, VII), with the distorted square pyramidal geometry. Though the exact order of events and level of protonation here is not known, a postulated catalytic cycle is shown. A lysine residue hydrogen bonded to an equatorial V-O oxygen activates the peroxide complex for  $X^-$  attack and a return to the



**Scheme 1** Haloperoxidase HOX generation and utilisation. (A and B) Biocatalytic cycle of heme iron-dependent and vanadium-dependent haloperoxidases, respectively, showing the release of freely diffusible hypohalous acid (NB the order of events in B remains unclear, with the protonation state of VII being ambiguous).<sup>21</sup> (C) Chlorination of 1,3-cyclopentanedione **12** by HOCl in the presence of a heme iron-dependent chloroperoxidase (CPO) and hydrogen peroxide suggested to occur in caldariomycin **14** biosynthesis.<sup>15</sup> (D) The reaction of monochlorodimedone (MCD) **15** with HOX, the benchmark assay for haloperoxidase activity.<sup>22</sup>





trigonal bipyramidal geometry (Scheme 1B, VIII) is observed. The resulting V–OX species undergoes hydrolysis to generate electrophilic  $X^+$  in the form of free HOX.<sup>26</sup> Electrophilic aromatic substitution occurs, as in the case with heme-iron haloperoxidases.<sup>27</sup> A notable difference with the heme iron-dependent biocatalytic cycle is that the vanadium centre maintains its oxidation state throughout the process.<sup>28</sup>

Haloperoxidases had initially been considered to lack significant enzyme-mediated selectivity due to the release of freely diffusible hypohalous acid. The free electrophilic species allows halogenation at the most electron-rich carbon, and a mixture of mono-, di- and tri-substituted products can often be observed. Detection of the free HOX through reaction with nucleophilic monochlorodimedone (MCD) **15** has previously been used to assay for haloperoxidase activity (Scheme 1D).<sup>29</sup> However, for a number of vanadium-dependent haloperoxidases, substrate, regio- and even stereo-specific halogenation is evident, as noted in the biosynthesis of napyradiomycins **76**,<sup>30</sup> and merochlorins **77**, **78**.<sup>31</sup> (Scheme 5). Currently, the exact mechanistic rationale for this specificity is not known. Here, free HOCl is notably not detected in assay with MCD.<sup>31</sup> It is thus postulated that residues within the enzyme might transiently support the electrophilic HOCl, whilst the electron-rich substrate is appropriately and proximally positioned. Moore invokes the involvement of an active site lysine forming a haloamine, in a manner analogous to the FDHs.<sup>31,32</sup> Though NapH1 from napyradiomycin biosynthesis is associated with the first chlorination event, this enzyme is also capable of mediating nonspecific bromination through generation and release of HOBr. Greater understanding of the mechanism enabling selective halogenation by these series of enzymes would be valuable, potentially enabling informed discovery of further enzymes with biotechnological potential.

**1.1.2 Flavin-dependent halogenases.** The haloperoxidases remained the only known halogenases for several decades until in 1995, the first of a second series of halogenases, the flavin-dependent halogenases (FDHs), Chl from 7-chlortetracycline **73** biosynthesis in *Streptomyces aureofaciens* was revealed in a study by Dai. The enzyme was implicated as being involved in chlorination with knockouts of the gene encoding Chl lacking halogenated product. However, from these initial studies, it was not fully apparent that this new series of enzymes were flavin-dependent, only a partial protein sequence lacking 100 amino acids at the N-terminal end including the flavin binding site. *In vitro* analysis of the tryptophan 7-halogenase PrnA, by van Pée and co-workers, demonstrated the powerful regioselectivity of this new class of halogenase and the requirement for both flavin and molecular di-oxygen.<sup>34,35</sup>

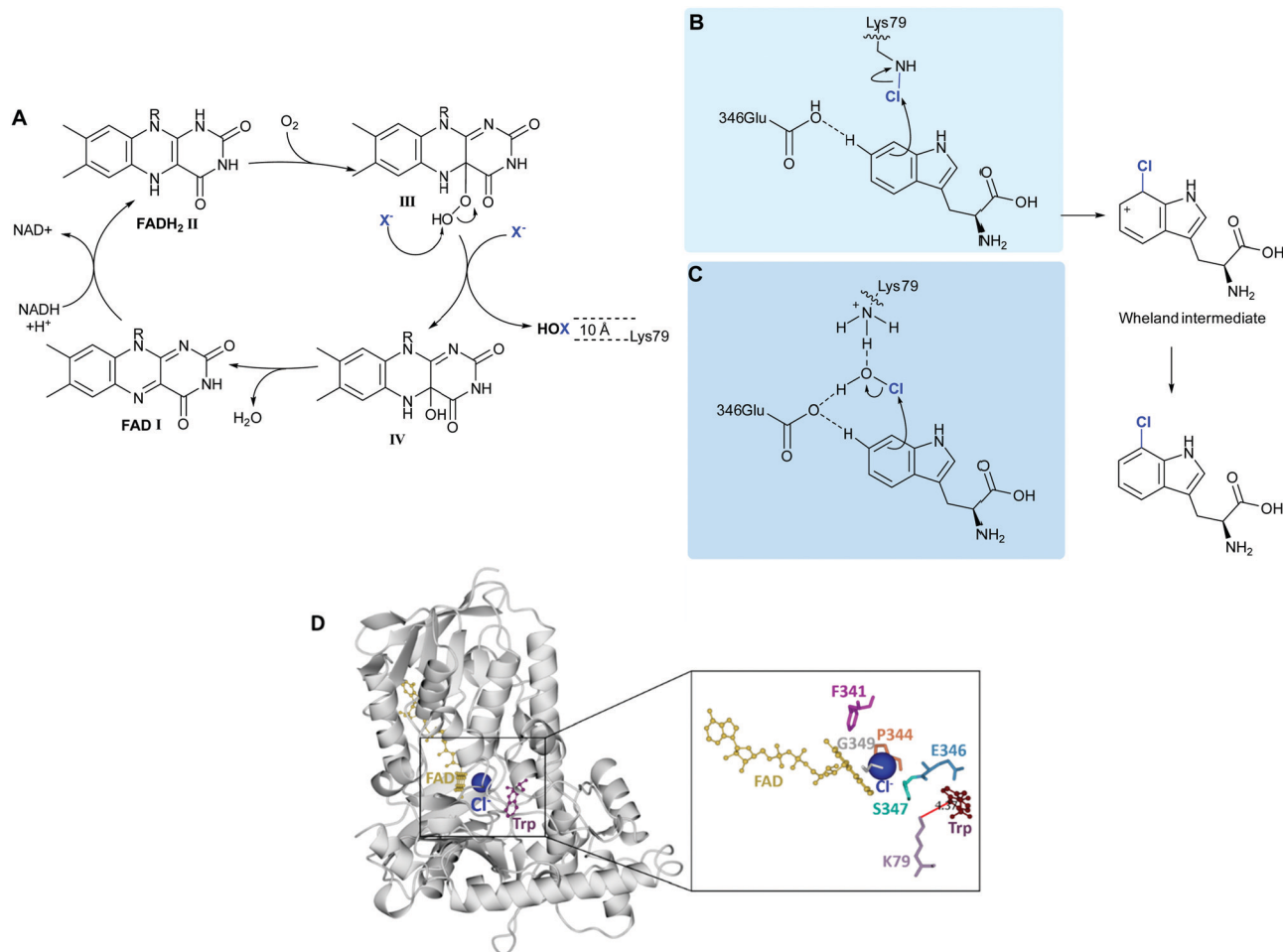
The FDHs can be separated into those acting on free substrates (variant A), and those requiring their substrate to be covalently tethered to a carrier protein *via* a phosphopantetheine linker (variant B).<sup>36</sup> The most extensively investigated and engineered FDHs are the tryptophan halogenases, variant A FDHs enabling regioselective halogenation of tryptophan at the 5, 6 or 7 positions. The first structural insights as to how

FDHs mediate their exquisite regioselectivity were revealed by van Pée and Naismith.<sup>37</sup> However, 15 years on and debate still remains as to the exact way in which these enzymes function. Details as to which residues within the active site participate may vary from one enzyme to another. The mechanisms proposed by Walsh<sup>38</sup> and van Pée<sup>39</sup> for the flavin dependent tryptophan halogenase are presented in Scheme 2B and C.

Reduced flavin adenine dinucleotide (FADH<sub>2</sub>), (Scheme 2A, I) and dioxygen in the presence of chloride, bromide or iodide are utilised by these enzymes to catalyse C–H to C–X. Freely diffusible FAD binds to a conserved GxGxxG motif. The reduced cofactor, FADH<sub>2</sub>, (Scheme 2A, II) reacts with molecular oxygen to form FAD(C4 $\alpha$ )–OOH (III), which subsequently reacts with a nucleophilic halide releasing HOX and FAD(C4 $\alpha$ )–OH (Scheme 2A, IV).<sup>40</sup> Unlike haloperoxidases, the HOX generation and electrophilic aromatic substitution (EAS) events are spatially separated by a  $\sim 10$  Å tunnel. At the end of this tunnel is a conserved lysine, proposed by Walsh to react with the HOX thereby forming a haloamine intermediate, guiding the regiochemistry of halogenation. (Scheme 2B).<sup>38</sup> Mutation of this lysine residue to alanine, across a series of halogenases including RebH, PrnA, and Rdc2, results in loss of activity.<sup>41</sup> In a subtly different mechanism, van Pée invokes a proximal glutamate residue, as well as the lysine residue binding to HOX, thereby activating it and positioning it (Scheme 2C).<sup>39</sup> In both models, the electrophilic halogenating species is directed towards the substrate, promoting a regioselective EAS reaction. The exact mechanism and halogenating species remain elusive. Mutation of this active site glutamate to glutamine in PrnA results in a drop in rate by two orders of magnitude.<sup>37</sup> This glutamate is not present in the more electron-rich phenol halogenases. A different proposed role for this active site glutamate, invoked in deprotonation of the Wheland intermediate indole benzo ring in the tryptophan halogenases, is that it perhaps plays a role in reducing the activation energy of halogenation of these less electron-rich species.<sup>41</sup> Notably no kinetic isotope effect is observed for halogenation of tryptophan by RebH.<sup>42</sup>

These FDHs share a similar flavin binding domain to the flavin-dependent monooxygenases, and though most utilise free and diffusible FAD, which is converted to FADH<sub>2</sub>, variants exist where the flavin is covalently bound to the enzyme, such as in the case of CmlS involved in chloramphenicol biosynthesis.<sup>44</sup> Naturally, most FDHs are usually provided the reduced flavin by a partner flavin reductase that utilises NADH (nicotinamide adenine dinucleotide) and FAD to generate reduced FADH<sub>2</sub>.<sup>35</sup> The partner flavin reductase can be readily substituted for any other capable of producing FADH<sub>2</sub>. Single component FDHs, that have this reductase function present within the one protein, also exist; the first discovered example of such an enzyme was Bmp5.<sup>45</sup> This enzyme shows other notable differences to all previously studied FDHs; in analogy to the decarboxylative hydroxylation observed for a single component monooxygenase, decarboxylative dibromination of **18** occurs (Fig. 2F).<sup>45</sup> Other intriguing variations of the FDHs also exist, recently, the first of a new class of zinc-binding FDHs, MaIA from the biosynthesis of the malbrancheamides **21** and **22** (Fig. 2G) was reported.





**Scheme 2** FDH-mediated electrophilic halogenation. (A) Biocatalytic cycle of a flavin-dependent halogenase, showing the release of hypohalous acid. HOX is channelled through a  $\sim 10$  Å tunnel. The regioselectivity of the electrophilic aromatic substitution is postulated to be achieved by: (B) a lysine residue, at the end of the tunnel that reacts with HOX to form a haloamine intermediate,<sup>43</sup> and (C) lysine and glutamine, acting in concert to position HOX and further electronically activate the tryptophan.<sup>39</sup> (D) Extrapolation of the crystal structure of PrnA showing how Lys79 is positioned over C7 of the indole ring mediating regioselective halogenation.

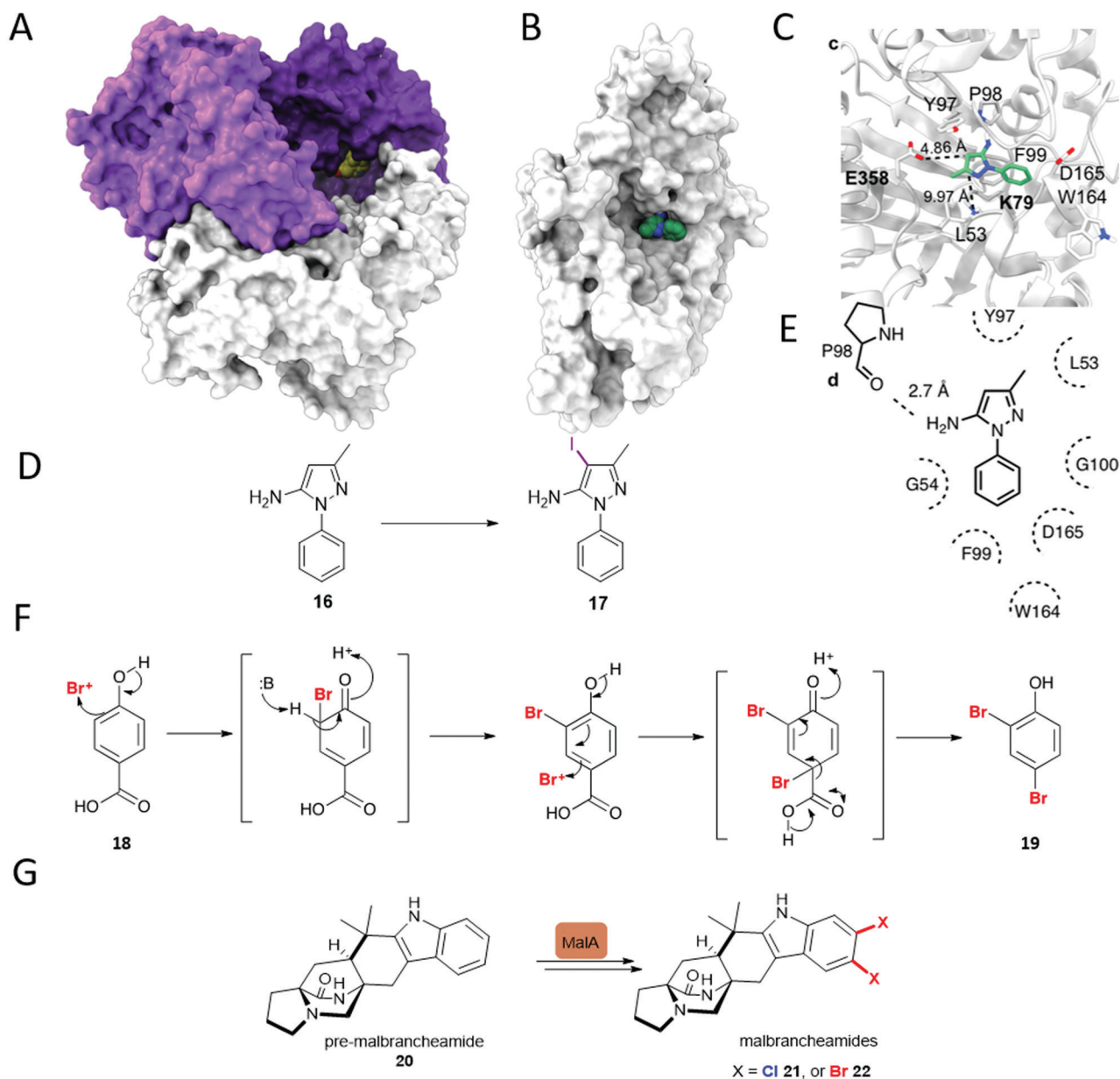
This remarkable enzyme is capable of mediating the late stage dichlorination of 5, and 6 positions of the indole benzo ring of pre-malbrancheamide **20**, a complex prenylated indole alkaloid. In the first step, positions 5 and 6 appear to be halogenated with comparable rates, perhaps due to increased electron density, the second halogenation even occurs far more quickly. Within the enzyme, the active site glutamate is seen proximal to the NH of the indole. It is clearly implicated in deprotonating this moiety, rendering the substrate more nucleophilic, and whilst mutation of this glutamate to aspartate results in low but observable activity, mutation to glutamine renders the enzyme inactive.<sup>46</sup>

FDHs contain multiple sequence motifs; a series of amino acid residues integral for enzyme activity or structure, that are generally conserved throughout all FDHs. These are a dinucleotide-binding motif GxGxxG, part of the Rossman fold allowing for binding of flavin and NAD(P)H,<sup>49</sup> and a tryptophan motif WxWxIP, suggested to block binding of substrate near the bound cofactor, thus preventing a monooxygenation reaction instead of halogenation.<sup>37</sup> It is notable though, that GxGxxG,

the flavin binding motif is present in almost all flavoenzymes. The WxWxIP motif or variants of this motif, is known to be absent in a number of FDHs including the more unusual Bmp5.<sup>48</sup> A new motif, Fx.Px.Sx.G was recently reported. This, or variants of this motif in which G is replaced with a charge bearing residue, line the tunnel between the flavin binding site and the site of halogenation, may be found in all active FDHs.<sup>47</sup> As such, it can be used as a method for quickly identifying candidate FDHs from uncured sequence data. A complementary and overlapping motif FX[DE]PX[EFL] was subsequently identified by Lewis and coworkers.<sup>157</sup>

Notably, almost all halogenases studied, until recently, have been identified from bacteria and fungi, most being found due to their implication in the biosynthesis of a specific natural product or series of natural products. Iodination had also been missing from the FDH portfolio of reactions, though postulated to be involved in the biosynthesis of iodinated natural products.<sup>50</sup> Whilst efforts had been made to explore the utilisation of enzymes such as RebH for iodination, incubation of such





**Fig. 2** Unusual FDHs. (A) Trimer of VirX1, the first FDH with a preference for iodination, and the first halogenase from a virus showing the large and accessible active site, with a bulky substrate **16** modelled in the active site. (B) View of a subunit of VirX1, again with substrate **16** modelled into the active site, and embedded in a hydrophobic cleft. (C) Key hydrogen bonds between VirX1 and **16** as predicted by structural modelling. (D) Regiospecific iodination of **16** by VirX1.<sup>47</sup> (E) Model of compound **16** in the active site of VirX1. (F) Decarboxylative bromination of **18** observed for Bmp5.<sup>48</sup> (G) Iterative dihalogenation in malbrancheamides **21** and **22** biosynthesis, by an unusual zinc-containing FDH.<sup>46</sup>

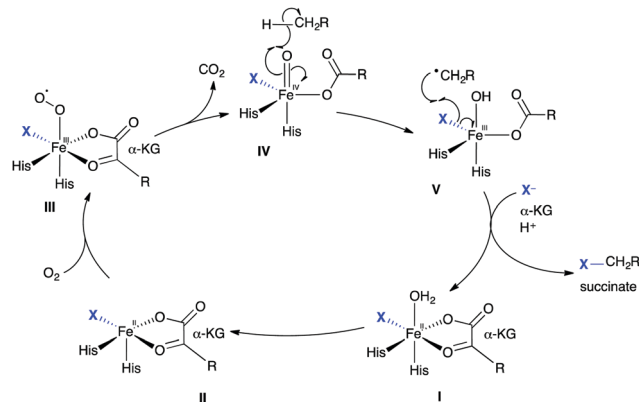
enzymes with iodide had been unfruitful, and competition studies had revealed that the presence of such salts precluded conversion, perhaps due to formation and inability to release the bulky product.<sup>43</sup> More recently the FDH, PltM, was shown to be able to process sodium iodide, though preference for bromination, then chlorination was apparent.<sup>51</sup> VirX1, the first halogenase to be characterised from a virus, is the first FDH for a preference for halogenation, and may be seen to have a large and readily accessible active site (Fig. 2A and B). Hydrophobic residues anchor the substrate in position, enabling regiospecific halogenation as guided by the proximal lysine (K79). Glutamate enables further

activation, of otherwise less electrophilic substrates. It is postulated that by virtue of openness of this active site not only can a large range of small to sterically bulky substrates be processed, but halogenation by the largest and most readily oxidised of the halides is enabled.<sup>47</sup>

**1.2 Radical (X<sup>•</sup>) halogenation.** The halogenation of less electron-rich alkyl groups requires a different enzymatic strategy compared to the halogenation of activated aryl moieties.<sup>52</sup> To achieve these energetically less favourable halogenations, radical chemistry is employed. Just as electrophilic FDH utilises similar enzymatic machinery to flavin dependent monooxygenases,







**Scheme 3** Radical halogenation. Biocatalytic cycle of a non-heme iron  $\alpha$ -ketoglutarate-dependent halogenase, showing the release of halogenated product.

analogously, non-heme iron  $\alpha$ -ketoglutarate (KG)-dependent (NHFe) halogenases (NHFeHals) may well have evolved from their hydroxylase counterparts. Understanding of the catalytic cycle mediated by the NHFeHals is informed by the studies of the hydroxylases, which also require a non-heme iron centre,  $\alpha$ -KG and  $O_2$  to function (Scheme 3). The catalytic cycle may notionally be considered to start with octahedral  $Fe^{II}$  coordinated in a bidentate fashion by  $\alpha$ -KG, and a weakly bound water ligand that is subsequently replaced by molecular oxygen. The remaining three positions, a “facial triad,” in hydroxylases, consists of two histidine residues and the carboxylate of either an aspartate or glutamate residue at the active site.<sup>53</sup> In halogenases, this carboxylate residue is usually found to be replaced by either an alanine or glycine, thereby enabling the halogen ion to coordinate at this vacant coordination site (Scheme 3, I). It has been shown that the weakly bound axial water ligand is initially displaced by substrate binding, triggering the subsequent reaction of II with molecular oxygen.<sup>54</sup> Oxidative decarboxylation of III affords trigonal bipyramidal IV, with succinate coordinated, and is the driving force for the catalysis. This high valent, high energy  $Fe^{IV}$ -oxido species is capable of abstracting  $H^\bullet$  from an otherwise unactivated alkyl group.<sup>55</sup> The regio-selectivity of abstraction is governed by how the substrate is positioned within the enzyme's active site. Selectivity for reaction of generated substrate radical with  $Fe(III)$ -X bond, leading to the desired halogenation over the  $Fe(III)$ -OH bond, which would lead to hydroxylation, is thought to be controlled by careful substrate positioning over  $Fe^{III}$  species V.<sup>56</sup> In HctB, a NHFeHal from hectochlorin biosynthesis, residues such as the Glu223 and Arg254 are thought to influence the charge density of the chlorine ligand, and are thus postulated to further promote halogenation.<sup>57</sup> The resultant alkyl radical proceeds to extract  $X^\bullet$  from V, returning the metal centre to its  $Fe^{II}$  state.

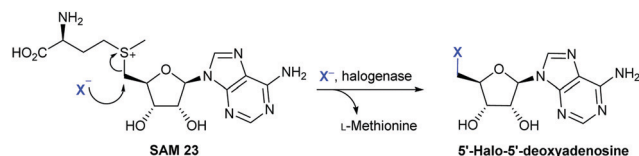
Significant conformational changes are observed within these NHFe proteins. These changes are postulated to occur directly after oxygen binding and the subsequent decarboxylation event, orientating the substrate for the observed regioselective halogenation. Crystal structures of WelO5 with NO (an  $O_2$  analogue) reveal a significant change in conformation. Initially, NHFeHals

appeared to be difficult to handle *in vitro* due to oxygen sensitivity of the  $Fe(II)$  core,<sup>58</sup> however, recent advances suggest an iron-reconstitution step can be performed to activate enzymes purified aerobically.<sup>59</sup> The majority of NHFeHals, such as SyrB1,<sup>58</sup> BarB1 and BarB2,<sup>60</sup> and CurA halogenase,<sup>61</sup> like the variant B FDHs, all required the substrate to be presented on carrier proteins, tethered *via* a phosphopantetheine linker (Ppant). In 2014, WelO5 was discovered, becoming the first member of a new family of NHFeHals (that we will refer to as variant A NHFeHals), which act on freestanding substrates.<sup>62</sup> Excitingly, these open up the possibility for using such proteins more readily in biotransformations.

**1.3 Nucleophilic ( $X^-$ ) halogenation.** Though the C-F bonds are prevalent in pharmaceuticals, and fluorine is the most abundant of the halogens on earth, few fluoro-metabolites are known. So far, 13 fluorometabolites, 7 of these being closely related fluorinated lipids, have been discovered. The bacterium *Streptomyces cattleya* is known to biosynthesise three of this limited portfolio of fluorinated metabolites: fluoroacetate, 4-fluorothreonine, and (2R,3S,4S)-trihydroxy-5-fluoropentanoic acid. These compounds are generated by the “fluorinase,” more formally named adenosyl fluoride synthase or FIA, that was discovered by O'Hagan and co-workers.<sup>63</sup> FIA is closely related to the more recently discovered SalL, an enzyme that mediates chlorination in salinosporamide biosynthesis.<sup>64</sup> These enzymes are members of a rare group of nucleophilic halogenases, all of which utilise *S*-adenosyl-L-methionine 23 (SAM) as substrate. *Streptomyces calvus* is capable of biosynthesising a fourth bacterial fluorometabolite, nucleocidin.<sup>65</sup> Intriguingly, a gene with similarity to that which encodes the fluorinase is absent from the genome of this organism implying that a further group of halogenases capable of mediating C-F bond formation exists. The dearth of fluoro-metabolites found in nature may perhaps be attributed to a number of factors. With the high electronegativity of fluorine, the possibility of oxidation to generate a fluorous electrophile or even a fluorous radical is energetically challenging, and no enzymes have so far shown this capability. Also, fluoride, though abundant, is generally present in minerals such as fluorospar (fluorite), which have high lattice energies and poor solubilities. Once in solution, fluoride is surrounded by a sheath of water, its high enthalpy of hydration ( $490 \text{ kJ mol}^{-1}$ ) precludes ready displacement of the water molecules to access the naked fluoride ion required for nucleophilic attack.<sup>66</sup>

The nucleophilic halogenases characterised so far all operate by holding the halide proximal to the electrophile, SAM. Crystallographic studies of the fluorinase reveal a small pocket, the fluoride binding site, giving insight into how displacement of the hydration sphere might be achieved. Whilst predominantly hydrophobic, this small pocket contains a serine.<sup>67</sup> Dehydration is proposed to occur when SAM binds with the positive charge on the sulphur possibly assisting with stabilisation of the naked fluoride. The requirement for the fluoride ion to bind first in order for a productive complex to be formed, is argued based on observed substrate inhibition by SAM. Within the crystal structure, SAM may be seen bound with its ribose ring forced





**Scheme 4** Nucleophilic halogenation. Halogenation of *S*-adenosyl-*L*-methionine **23** (SAM) by X<sup>-</sup> mediated by a nucleophilic halogenase, yielding *L*-methionine and 5'-halo-5'-deoxyadenosine.<sup>67</sup>

into a high energy, eclipsed conformation. The halide anion is positioned collinearly with the SAM C-S bond, thus promoting an S<sub>N</sub>2 backside attack (Scheme 4).<sup>67</sup> When the enzyme is crystallised in the presence of both SAM and fluoride, the products *L*-methionine and 5'-fluoro-5'-deoxyadenosine (5'-FDA) with the ribose ring in the relaxed conformation may be observed.

The fluorinase has a strong preference for fluoride, but can incorporate chloride to a low level.<sup>68</sup> SalL, from *Salinispora tropica* bacterium has 54% sequence similarity to the fluorinase, and in the halide binding pocket, a glycine residue resides in place of a serine. This enzyme shows a strong preference for chlorination and, whilst capable of utilising bromide and iodide, it cannot process fluoride.<sup>64</sup>

Notable uses of the fluorinase include exploration of its application to <sup>18</sup>F incorporation for positron emission tomography (PET),<sup>69</sup> and it's the utilisation of the fluorinase in parallel with a polyketide synthase to enable synthetic biological access to simple fluorinated polyketides.<sup>70</sup>

**1.4 Biocatalytic potential.** The halogenases offer great potential for use as biocatalysts. To realise the application of these important enzymes, the regio-control, -selectivity and reaction rate afforded by the different classes of enzymes is important to consider. The reported exploration of kinetics for the halogenases in the literature is both sparse and varied, and can be challenging to compare due to the different ways in which the data are obtained. We compile a representative overview of the kinetic analyses available for different families of halogenases for a brief comparison between families (Table 1), as well as a glossary of the halogenases that have been structurally characterised so far (Table 2). Haloperoxidases, such as CPO, may be seen to have high turn-over numbers

( $K_{\text{cat}}$  33 ± 1 min<sup>-1</sup>), yet such catalysts that may be utilised to halogenate a broad series of substrates (see Section 2) lack regio-control. Conversely, the NHFeHals also have a high turn-over number, mediate a highly regio-selective reaction, but the natural reaction that it catalyses is much slower than CPO ( $K_{\text{cat}}/K_{\text{M}}$  0.0525 vs. 1.22 μM<sup>-1</sup> min<sup>-1</sup>, Table 1).

## 2. Natural and unnatural aromatic substrate scope for biocatalytic halogenation

### 2.1 Indolic substrates

#### 2.1.1 Early stage halogenation in biosynthetic pathways.

Indoles are commonly utilised motifs in medicine and agrochemistry, and new tools that enable their chemical diversification are desirable. There are a series of FDHs that allow for this diversification, *via* halogenation. Many FDHs mediating the regiospecific halogenation of the indole moiety of tryptophan have been found so far. Most of these enzymes have been found to be involved in the early stages of bioactive natural product synthesis. Such enzymes include PrnA and RebH, tryptophan 7-halogenases, from the first step of the biosynthesis of the antibiotic pyrrolnitrin **24** and the indolocarbazole anticancer agent rebeccamycin **36**, respectively.<sup>35,43</sup> Tryptophan 6-halogenases include ThdH (alternatively known as ThaI) from the first step of thienodolin **25** biosynthesis, a plant growth promoting compound,<sup>71</sup> SttH identified *in silico* adjacent to an NRPS cluster,<sup>72</sup> SatH from *Streptomyces albus*,<sup>73</sup> BorH, a thermophilic halogenase from borregomycin **33–35** biosynthesis,<sup>74</sup> and AfnX responsible for generating the 6-chlorotryptophan precursor of alboflavusin A **26**. Alboflavusin is an analogue of the dimeric non-ribosomal peptide (NRP) anti-tumor agent and apoptosis inducing agents, himastatin and chloptosin.<sup>75</sup>

Though less common so far, 5-chlorotryptophans may also be seen as motifs in natural products. Examples include the antibiotic pyrroindomycin B **27**; the first step in the biosynthesis of this compound is the generation of 5-chlorotryptophan by PyrH.<sup>76</sup> Sequence analysis of *Xenorhabdus szentirmaii* revealed a further flavin dependent tryptophan 5-halogenase XszenFHal. Though the natural function of this halogenase remains unknown, it was demonstrated to be capable of

**Table 1** Representative electrophilic (VirX1, PltM, PyrH, CPO, V-BrHPO), nucleophilic (FIA) and radical halogenases (WelO5, PfHalA) and their kinetic parameters with respect to substrate at set concentrations of other cofactors. ND = no data published

Enzyme	Substrate	Classification	Halide	$k_{\text{cat}}$ (min <sup>-1</sup> )	$K_{\text{M}}$ (μM)	$k_{\text{cat}}/K_{\text{M}}$ (μM <sup>-1</sup> min <sup>-1</sup> )	Ref.
VirX1	6-Azaindole	FDH	I	5 ± 0.5	28 ± 2	0.179 ± 0.004	47
PltM	Resorcinol	FDH	Br	2.4 ± 0.6	53 ± 3	0.044 ± 0.008	51
			Br	1.9 ± 0.1	0.71 ± 0.3	2.7 ± 0.6	
			Cl	2.3 ± 0.1	0.076 ± 0.013	30 ± 5	
PyrH	Tryptophan	FDH	Cl	3.56 ± 1.1	109 ± 44	0.0325 ± 0.017	51
KtzQ	7-Cl-L-Tryptophan	FDH	Cl	1.4	114	0.012	82
MalA	Premalbrancheamide	FDH	Cl	0.08 ± 0.05	7.0 ± 2.9	11.49 ± 0.02	46
WelO5	12-Epi-fisherindole U	NHFeHal	Cl	1.8 ± 0.2	ND	ND	83
PfHalA	Lysine	NHFeHal	Cl	14.7 ± 0.71	280 ± 60	0.0525 ± 0.0118	84
AdeV	2'-Deoxyadenosine monophosphate	NHFeHal	Cl	0.6	2500	0.24 × 10 <sup>-3</sup>	85
CPO	Catechol	HPO	Cl	33 ± 1	27 ± 8	1.22 ± 0.36	86
V-BrHPO	2-Chlorodimedone	V-HPO	Br	5	ND	ND	26
FIA	SAM	Nucleophilic	F	0.07 ± 0.001	6.5 ± 0.3	0.0108 ± 0.0005	87





Table 2 Glossary of structurally characterised halogenases, arranged by class

Enzymes		No. residues	PDB code	Cofactors and ligands	Identification time	Relevant BGC	Oligomer	Reported substrates	Halogen substrates	Halogen stage	Source	Subgroup	Ref.
PnA	538	2AR8		FAD, Cl <sup>-</sup> , 7-Cl-Trp	2005	Pyrrolnitrin	Dimer	Tryptophan 7C, indolic substrates	Cl, Br	Early stage	<i>Pseudomonas fluorescens</i>	A	37
		2APG		FAD, Cl <sup>-</sup>									37
		2AQJ		FAD, Cl <sup>-</sup> , Trp									37
		2ARD		FDA									37
		4Z43 (E450K)		FAD, Cl <sup>-</sup>	2015								79
		4Z44 (F454K)		FAD, Cl <sup>-</sup> , PO4									79
RebH	530	2OAI		FAD, ADN, Trp, Cl <sup>-</sup>	2008	Rebeccamycin	Dimer	Tryptophan 7C, indolic substrates, aniline substrates	Cl, Br	Early stage	<i>Lechevalieria aerocolonigenes</i>	A	88
		2O9Z		PO4									88
		2OAL		FAD, Cl <sup>-</sup>	2007								43
		2OAM		Apo form									43
		2E4G		Trp									43
		4LU6 (thermostabilized)		PO4	2013								89
CndH <sup>a</sup>	512	3E1T		FAD	2020								TBP
				6P00 (11 mutations)									TBP
				6P2V (6 mutations)									TBP
				7JU0 (1 mutation)									TBP
PyrH	511	2WET		FAD, Trp, Cl <sup>-</sup>	2010	Chondrochloren	Monomer	Aliphatic substrates	Cl	Early stage	<i>Chondromyces crocatus</i>	B	36
		2WEU		Trp	2009	Pyrroindomycin	Dimer	Tryptophan 5C	Cl, Br	Early stage	<i>Streptomyces rugosporus</i>	A	90
		2WES (E46Q)		FAD, Cl <sup>-</sup>	2009								90
CmlS <sup>a</sup>	570	3I3L		FAD	2010	Chloramphenicol	Monomer	Aliphatic substrates	Cl, Br	Late stage	<i>Streptomyces venezuelae</i>	A	44
PltA	455	5DBJ		FAD, Cl <sup>-</sup>	2015	Pyoluteorin	Dimer	Pyrrolic substrates	Cl	Early stage	<i>Pseudomonas fluorescens</i>	B	91
Mpy16	447	5BUK		FAD	2016	Pyrrolyl-S-Mpy15	Monomer	Pyrrolic substrates	Cl	Late stage	<i>Streptomyces</i> sp. CNQ-418	B	48
Th-Hal	510	5LV9		—	2016	—	Monomer	Tryptophan	Cl, Br	—	<i>Streptomyces violaceusniger</i>	A	92
SttH	519	5LVA 5HV5		FMN FAD, Cl <sup>-</sup>	2016 2016	—	Dimer	Tryptophan 6C, N-methyltryptophan, kynurenine, anthranilamide	Cl, Br	—	<i>Streptomyces toxytricini</i>	A	80
Bmp2	409	5BVA		FAD, EDO	2016	Tetrabromopyrrole	Monomer	Pyrrolic substrates	Br, I	—	<i>Pseudoalteromonas luteoviolacea</i>	B	48





Table 2 (continued)

No. Enzymes	residues	PDB code	Cofactors and ligands	Identification time	Relevant BGC	Oligomer	Reported substrates	Halogen stage	Halogenation	Source	Subgroup Ref.
MalA	667	5BUL (Y302S F306V A345W)	FAD	2017	Malbranchamide, malbranchamide B, isomalbranchamide B	Monomer	Premalbranchamide, malbranchamide B, isomalbranchamide B	Cl, Br	Late stage	<i>Malbranchea aurantiaca</i>	A
		5WGX (H253A)	FAD, IM7, Cl <sup>-</sup>	2017							
		5WGV (C112S C128S)	FAD, PM7, Cl <sup>-</sup>	2017							
		5WGU (E494D)	FAD, PM7, Cl <sup>-</sup>	2017							
		5WGX (H253A)	FAD, MB5, Cl <sup>-</sup>	2017							
		5WGY (C112S C128S)	FAD, MB5, Cl <sup>-</sup>	2017							
		5WGR	FAD, PM7, Cl <sup>-</sup>	2017							
		5WGS (H253F)	FAD, PM7, Cl <sup>-</sup>	2017							
		5WGT	FAD, PM7, Cl <sup>-</sup>	2017							
		5WGW	FAD, MB5, Cl <sup>-</sup>	2017							
BrvH	502	6FRL	—	2018	—	Dimer	Indolic substrates	Br > Cl	—	<i>Brevundimonas BAL3</i>	B
Tar14	532	6NSD	FAD	2019	Taromycin	Dimer	Tryptophan	Cl, Br	Early stage	<i>Saccharomonospora</i> sp. CNQ-490	A
Thal	534	6H43	PO4	2019	Thienodolin	Dimer	Tryptophan 6C, indolic substrates	Cl, Br	Early stage	<i>Streptomyces albogriseolus</i>	A
PltM	522	6H44	Trp, PO4	2019	Pyoluteorin	Monomer	Phenolic substrates, aniline substrates	Cl, Br, I	—	<i>Pseudomonas protegens Pf-5</i>	A
		6IB5 (Thal-RebH5, 5 mutations)	Trp, PO4	2019							
		6SLT	FAD, AMP, PO4	2019							
		6SLS	Trp, PO4	2019							
		7AQV (N-6His-Thal-RebH5, 5 mutations)	FAD, PO4	2020							
		7AQV (N-6His-Thal-RebH5, 5 mutations)	—	2020							
		7AQV (N-6His-Thal-RebH5, 5 mutations)	BCN, SER, ALA, GLY	2020							
		7CU0	Trp	2020							
		7CU1	FAD, AMP	2020							
		7CU2	FDA	2020							
PltM	522	6BZN	—	2019	Pyoluteorin	Monomer	Phenolic substrates, aniline substrates	Cl, Br, I	—	<i>Pseudomonas protegens Pf-5</i>	A
		6BZZ	FAD	2019							



Table 2 (continued)

Enzymes residues	No.	PDB code	Cofactors and ligands	Identification time	Relevant BGC	Oligomer	Reported substrates	Halogenation		Subgroup Ref.	
								Halogen stage	Source		
Virx1	531	6QGM	Apo form	2019	—	Trimer	Aromatic substrates, heterocycles, azaspirocycles	I > Br > Cl	—	47	
		6BZA	FAD, 13X, Cl <sup>−</sup>	2019						51	
		6BZQ	FAD, Br <sup>−</sup> , Cl <sup>−</sup>	2019						51	
		6BZT (L111Y)	FAD, Br <sup>−</sup> , Cl <sup>−</sup>	2019						51	
		6BZI	EMC, HG	2019						51	
<hr/>											
Classification	No.	Enzymes residues <sup>b</sup>	PDB code	ligands	Cofactors and Idt. time	Relevant BGC	Oligomer	Reported substrates	Halogen Source	Ref.	
Vanadium-dependent haloperoxidases		Ci-VCPO	609	1VNC	VO4, AZI	1996	—	Monomer	Chlorinated 1-(4-ethoxy-3-methoxyphenyl)-2-(2-methoxyphenoxy)-1,3-dihydroxypropane	Cl, Br <i>Curvularia inaequalis</i>	97
				1VNS	SO4	1999					98
				1VNI	Apo form						98
				1VNH	VO4						98
				(H496A)							
				1VNG	VO4						98
				(H404A)							
				1VNF	VO4		—				98
				(R360A)							
				1VNE	VO4						98
				(D292A)							
				1IDQ	VO4	1997	—				99
				1IDU	VO4	—					99
				3BB0	PO3	2008					100
		An-VBPO1	556	1QI9	VO4, I <sup>−</sup>	1999	—	Dimer	2-Methylindole, 2-phenylindole, phenolsulfonephthalein, derivatives of O-methyl pyrrole-2-carboxylate	Br, I <i>Ascophyllum nodosum</i>	101–103
		An-VBPO2	597	5AA6	VO4	2015	—	Hexamer	Refer to An-VBPO1	Br, I <i>Ascophyllum nodosum</i>	TBP
		Co-VBPO	596	1QHB	PO4	1999	—	Dodecamer	( <i>E</i> )-(+)-Nerolidol	Br <i>Corallina officinalis</i>	104
		Cp-VBPO	598	1UP8	PO4	2005	—	Dodecamer	<sup>c</sup>	Br <i>Corallina pilulifera</i>	105
		NapH1	531	3W35 (Apo form)	—	2012	Napyradiomycin	Dimer	Polyketide-terpenoid substrates	Cl, Br <i>Streptomyces</i> sp. C1NQ525	TBP
		Zg-VIPO1	458	4CIT	VO4	2012	—	Monomer	TB iodoperoxidase assay	I <i>Zobellia galactanivorans</i>	TBP
					2014	—				106	

Table 2 (continued)

Classification	Enzymes	No. residues <sup>b</sup>	PDB code	Cofactors and Idt. time	Relevant BGC	Oligomer	Reported substrates	Halogen Source	Ref.
Heme-dependent haloperoxidases	Zg-VIPO2	458	4USZ	VO4	2014 —	Monomer	TB iodoperoxidase assay	I	106
	AmVHPO	639	5LPC	PO4	2016 —	Dodecamer	Aromatic substrates	Cl, Br	107
	HL-HPO	373	1CPO	HEM	1995 Caldariomycin	Monomer	1,3-Cyclopentanedione, electron-rich aromatic substrates	Cl, Br	108
			2CPO	HEM					108
			2CJO	HEM	2006				109
			2CIW	HEM, I <sup>−</sup>					109
			2CIV	HEM, Br <sup>−</sup>					109
			2CJ2	HEM, FMT					109
			2CJ1	HEM, FMT					109
			2CIZ	HEM, Br <sup>−</sup>					109
NHF dependent halogenases (NHFeHal)			2CIY	HEM, Br <sup>−</sup> , CYN, DMSO					109
			2CIX	HEM, CEJ, Br <sup>−</sup> , EDO					109
	CytC3	319	3GJA	ACT	2009 $\gamma, \gamma'$ -Dichloroaminobutyrate	Monomer	L-2-Aminobutyric acid	Cl	110
			3GJB	AKG, ACT, Fe <sup>2+</sup>					110
	SyrB2	310	2FCT	DSU, AKG, Cl <sup>−</sup> , Fe <sup>2+</sup>	2006 Syringomycin E	Monomer	L-Threonine	Cl, Br	111
			2FCV	DSU, AKG, Br <sup>−</sup> , Fe <sup>2+</sup>					111
			2FCU	DSU, AKG					111
	CurA-Hal	2311	2LIU	—	2011 Curacin A	Domain (NMR structure)	3-Hydroxyl-3-methyl-glutaryl	Cl	111
			2LIW	PNS, MAH	2011	Domain (NMR structure)			111
	WelO5	315	5IQS	AKG, Fe <sup>2+</sup> , Cl <sup>−</sup>	2016 Welwitindoline	Monomer	12-Epi-fischerindole U	Cl, Br	111
			5IQT	6CU, AKG, Fe <sup>2+</sup> , Cl <sup>−</sup>					111
			5IQU (G166D)	6CU, AKG, Fe <sup>2+</sup>					111
			5IQV	6CU, AKG, Fe <sup>2+</sup> , Cl <sup>−</sup> , NO					111
			5J4R	AKG, Ni <sup>2+</sup>	2016				TBP
			5T22	AKG, Ni <sup>2+</sup>	2016				TBP
			5TRQ	SIN, ACT, Ni <sup>2+</sup>	2016				TBP





Table 2 (continued)

Classification	Enzymes	No. residues <sup>b</sup>	PDB code	Cofactors and Idt. time	Relevant BGC	Oligomer	Reported substrates	Halogen Source	Ref.
SAM (S-adenosyl-L-methionine)-dependent halogenases	FIA	299	1RQP	SAM	2004 <sup>d</sup>	Hexamer	SAM	F	67
			1RQR	5FD, MET				<i>Streptomyces catleya</i>	67
			2C2W	5CD, Cl <sup>-</sup>	2005				68
			2C4T	SA8, Cl <sup>-</sup>	2007				TBP
			2C4U	—	2006				112
			2C5B	5F1, MET					112
			2CBX	CC5					112
			2CC2	—					112
			2C5H	3D1, MET, Cl <sup>-</sup>	2006				TBP
			2V7T	SAH, Cl <sup>-</sup>	2008				87
			(S158G)						87
			2V7U	SAM, Cl <sup>-</sup>					87
			(S158G)						87
			2V7V	5FD					87
			2V7W	5FD					87
			(S158G)						87
			2V7X	5FD, MET					87
			(S158G)						87
			4CQJ	EFA	2014				TBP
			5FIU	Y3J, TLA	2015				113
			5LMZ	1DA, Cl <sup>-</sup>	2016				114

<sup>a</sup> No *in vitro* activity demonstrated to date. <sup>b</sup> Natural protein length from UniProt. <sup>c</sup> Not specifically mentioned, tested by standard substrate monochlorodimedone/TB iodoperoxidase assay. <sup>d</sup> FIA is not in a BGC, but participates in the synthesis of fluoroacetate and 4-fluorothreonine. A. Flavin dependent halogenases. TBP To be published. TBP: To be published (compound structures reported in the PDP but not yet formally published elsewhere).



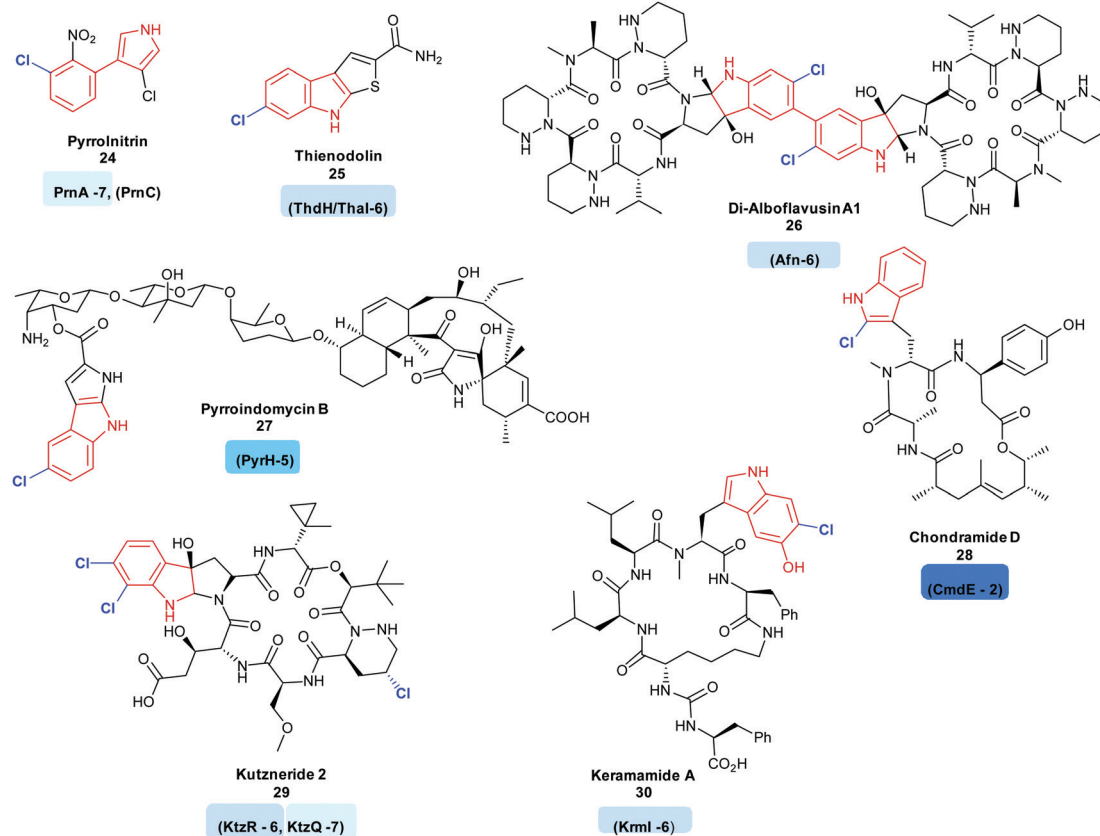
regioselectively halogenating a series of indolic and anthranilic substrates.<sup>77</sup>

The first successful demonstration of the ability to modify the regioselectivity of a FDH was *via* site directed mutagenesis of tryptophan 7-halogenase PrnA. The modification of a large phenylalanine residue in the active site to a smaller alanine allowed for a different binding orientation of tryptophan, leading to halogenation at the 5 position, though it must be noted that halogenation at the 7 position still dominated (2:1 ratio of 7-/5-bromination).<sup>78</sup> Despite useful changes in selectivity being rare, there are a number of further successes.<sup>42,79</sup> Positioning of the lysine residue over the substrate was thought to be critical to site of halogenation. Various studies toward gaining deeper understanding of subtle factors beyond the positioning of the lysine residue in the active site have been carried out, including structural comparison of PyrH, SttH and PrnA, and, specifically, a loop proximal to the active site, informing the design of mutants with modified regiochemistries.<sup>80</sup> Modifications to ThaI, to promote chlorination at the 7 position, rather than the 6 position have also been reported.<sup>81</sup>

Biosynthesis of 4-chlorotryptophan remains a mystery. Although it was isolated from immature pea seedlings in 1970,

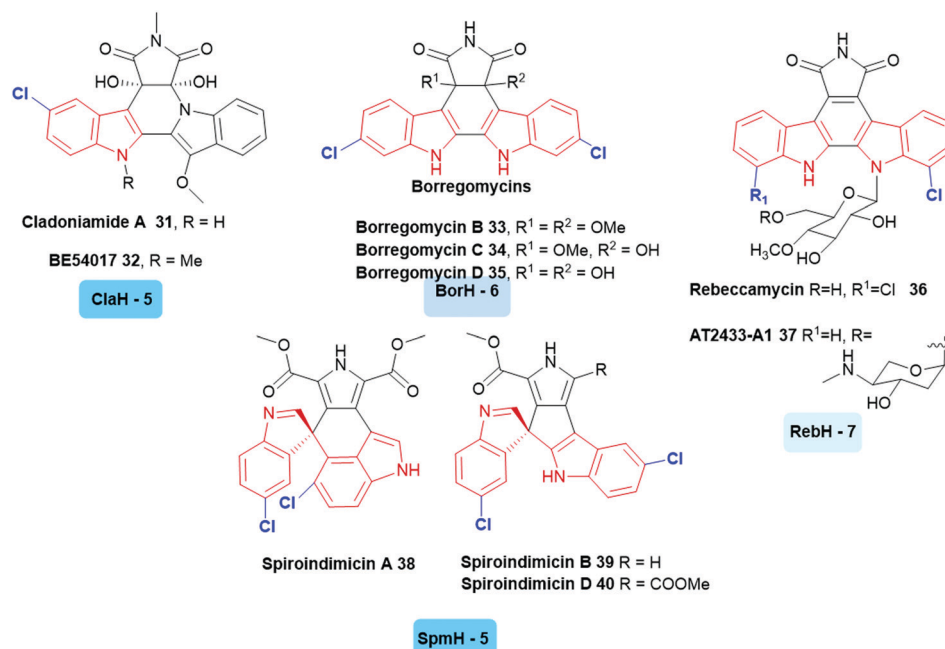
and postulated to be a precursor to a plant growth hormone,<sup>115</sup> the halogenase has not yet been found, hinting that it may be strikingly different to the other, well known FDHs. There have also been fewer reports of natural products containing 2-halotryptophans, perhaps due to their inherent instability. The keramamides **30**<sup>116</sup> and linked family of jaspamides, chondramides **28** are two exceptions, with CmdE implicated in the generation of 2-chlorotryptophan,<sup>117</sup> though for this enzyme, no activity with free tryptophan can be shown. It is not clear as to whether lack of observation of production of this species by CmdE is due to the enzyme requiring a different substrate or due to the instability of the product. 2-Chlorotryptophans are very unstable and both the free amine and carboxy groups might be considered to promote dechlorination.

Tryptophan halogenases capable of halogenating substituted tryptophans are also known. These include the tryptophan 6-halogenase KtzR, which acts in tandem with KtzQ in the first steps of kutzneride **29** assembly, installing a second chlorine into the 7-chlorotryptophan that KtzQ generates<sup>118</sup> and KrmI, another tryptophan-6-halogenase, which processes 4-hydroxy-tryptophan as its natural substrate, though the wild type enzyme shows ability to process a fairly broad range of



**Fig. 3** Halogenated natural products arising from the halogenation of indolic substrates: all include regioselective halogenation of tryptophan as a first step in their biosynthesis (tryptophan halogenase name and regiochemistry noted below each corresponding natural product). Examples include (C7) pyrrolnitrin **24**,<sup>120</sup> (C6) thienodolin **25**,<sup>71</sup> Di-alboflavusin A1 **26**,<sup>75</sup> (C5) pyrroindomycin B **27**,<sup>76</sup> (C2) chondramide D **28**,<sup>117</sup> (C6, C7) kutzneride **29**,<sup>121</sup> keramide A **30**.<sup>116</sup> Tryptophan halogenases involved in the biosyntheses of these natural products are depicted, and are colour coded in accordance with the regiochemistry of the halogenation that they mediate. Chlorination is indicated in blue. Different regio-chemistries for chlorination are represented by different shades of blue.





**Fig. 4** Naturally occurring halogenated bisindole alkaloids. These compounds incorporate a halogen into tryptophan in the first stage of their biosynthesis. (The tryptophan halogenase responsible and its regioselectivity that it confers, are noted below each natural product.) Examples include indolocarbazoles (C5) cladoniamide A **31**,<sup>122</sup> (C5) BE-54017 **32**,<sup>128</sup> (C6) borregomycins B **33**, C **34** and D **35**,<sup>123</sup> (C7) rebeccamycin **36**,<sup>40</sup> (C7) AT2433-A1 **37**,<sup>129</sup> (C5) spiroindimicins **38–40**.<sup>125</sup> Non-halogenated analogues of the latter have been accessed through deactivation of SpmH.<sup>125</sup> Tryptophan halogenases involved in the biosyntheses of these natural products are depicted, and are colour coded in accordance with the regiochemistry of the halogenation that they mediate. Chlorination is indicated in blue. Different regio-chemistries for chlorination are represented by different shades of blue.

other substrates. KrmI is an unusual halogenase, involved in the biosynthesis of the keramamides **30**, this FDH is translationally fused to a ThiF protein<sup>119</sup> (Fig. 3).

A large variety of symmetrically and asymmetrically halogenated indolocarbazole compounds are known, including the cladoniamides **31**, **32**,<sup>122</sup> borregomycins **33–35**,<sup>123</sup> and rebeccamycins **36**, **37**.<sup>40</sup> From genomic analysis of the producing organisms, the first step in the biosynthesis of these compounds also involved the generation of 5, 6 or 7-halotryptophan. This biosynthetic logic has been utilised to enable the generation of combinatorial libraries of known and novel compounds in which halogenation and glycosylation patterns are modulated.<sup>124</sup> Similar to these compounds are the bisindole alkaloids indimicins and related spiroindimicins **38–40** (Fig. 4).<sup>125</sup>

The tryptophan halogenases have been previously thought to possess only modest innate substrate flexibility. RebH is capable of processing tryptolines, whilst KrmI will halogenate unsubstituted indoles and fluoro-tryptophan.<sup>42,119</sup> The tryptophan halogenases continue to be subject to extensive rational reengineering and directed evolution due to their potential as tools for biocatalysis. In addition to studies that explore and modulate factors governing regiocontrol, considerable effort continues to be invested in extending substrate scope.<sup>41</sup> Wild-type RebH (the tryptophan 7-halogenase from rebeccamycin **36** biosynthesis) has been demonstrated to be capable of regioselectively halogenating a series of indolic (Fig. 5) and arene substrates (Fig. 9).<sup>126</sup>

Natural substrate scope for the FDHs RebH, GsfI, ThaI and Rdc2 is seen to be a little broader, and these enzymes have

capability to process a range of indolic substrates.<sup>42</sup> Directed evolution of RebH has been employed by the Lewis group to tune the enzyme for use in 7- **44**, 6- and 5-chlototryptamine generation enabling quantitative conversions in some instances, and high selectivities.<sup>42</sup> Evolved RebH variants have been employed, achieving selective late stage functionalisation of chlorination of sterically bulky substrates, including tricyclic tryptolines **50–53** and large indole carbazoles **56**, **57**, improving conversions of these unnatural substrates as well considerably extending substrate scope from what was previously known (Fig. 4).<sup>126,127</sup>

Production of halogenated indolic substrates by flavin-dependent halogenases is commonly reported at small scale.<sup>47,126,127</sup> The use of Cross-Linked Enzyme Aggregates (CLEAs), in which the physical aggregation of enzymes with a cross-linking agent is mediated, can enable enzyme stabilisation and enhanced yields. By crosslinking RebH and PrnF (the PrnA-related reductase) with glutaraldehyde, Sewald was able to achieve yields on a gram scale with CLEAs for the halogenation of L-tryptophan, D-tryptophan and L-5-hydroxytryptophan, establishing the foundations for halogenase biocatalysis on a preparative scale<sup>130</sup> (see Section 3).

**2.2 Phenolic substrates.** Phenolic quinolines attract attention for their many and varied medicinal properties such as antimalarial, antibacterial and anticancer activities.<sup>131</sup> Structure–activity relationship (SAR) reveal that installation of a chlorine or bromine atom within the quinoline ring of antitubercular quinolines can enhance their activity.<sup>131</sup> Beyond antitubercular uses, chlorine-





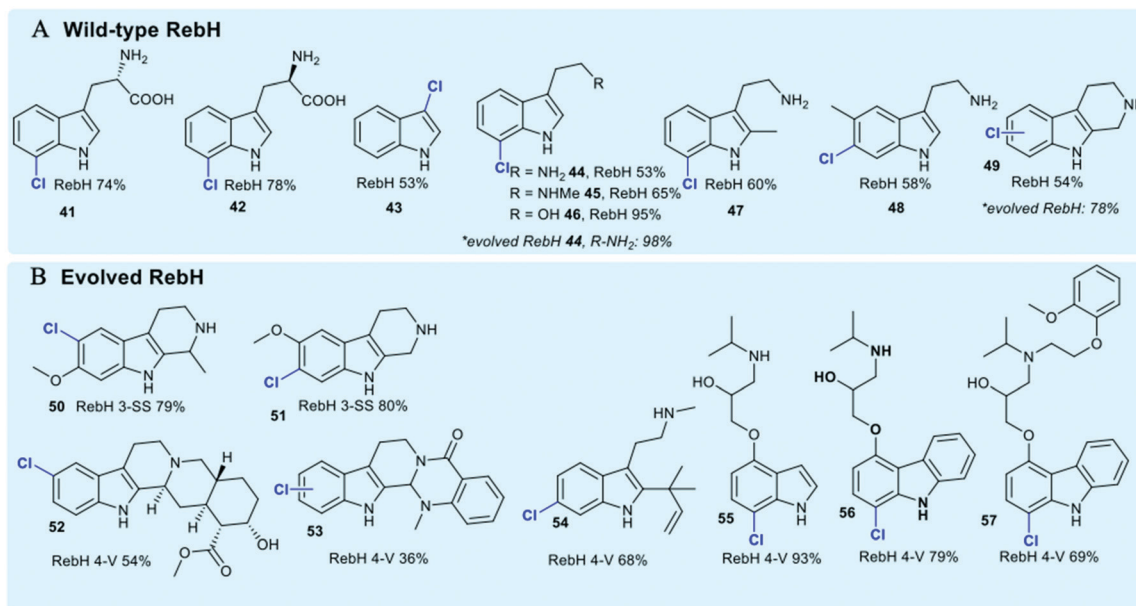
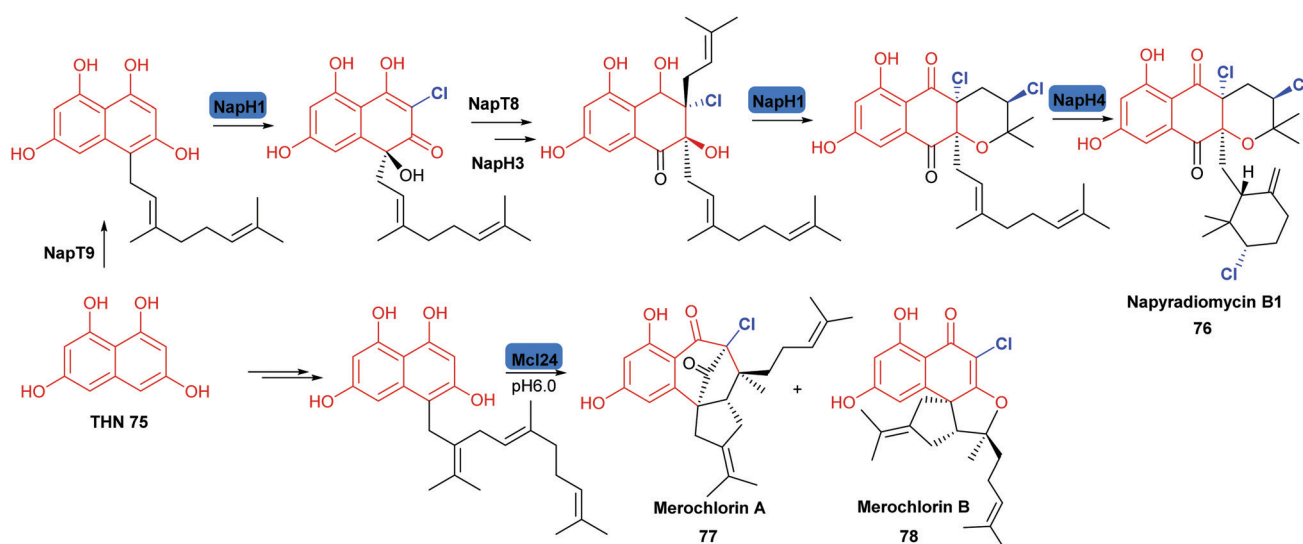


Fig. 5 Representation of some of the diverse substrate scope of tryptophan halogenases. Enzymes include (A) wild-type RebH chlorinating indoles;<sup>42,126</sup> (B) evolved RebH 3-SS and RebH 4-V which accept tricyclic tryptoline and large indole carbazoles.<sup>127</sup>

containing 8-hydroxyquinoline derivatives are being explored for treatment of Alzheimer's disease.<sup>132</sup> Phenols, being electron-rich, are frequently seen to be good substrates for electrophilic enzymatic halogenation (Fig. 8 and Scheme 5). While the haloperoxidases generally confer limited specificity, the FDHs show good levels of regiocontrol for accessing such halogenated compounds.

**2.2.1 Early stage halogenation by FDH variant B enzymes.** Whilst there has been considerable and detailed investigations of a number of the variant A halogenases that work on substrates that are not covalently tethered to a carrier enzyme

(in particular the tryptophan halogenases), fewer *in vitro* investigations of variant B halogenases have been published. However, a diverse range of bioactive natural products that contain halogenated phenols have been isolated. Here, the operation of variant B FDHs may be seen to be prevalent, and investigations indeed indicate many of the substrates to be enzyme tethered. Halophenol-containing natural products include the actinobacteria-produced aminocoumarins clorobiocin **58** and simocyclinone D8 **59**. The initial steps of the biosyntheses of these metabolites are postulated to include halogenation of enzyme-tethered tyrosine (mediated by Clo-hal and SimX1,

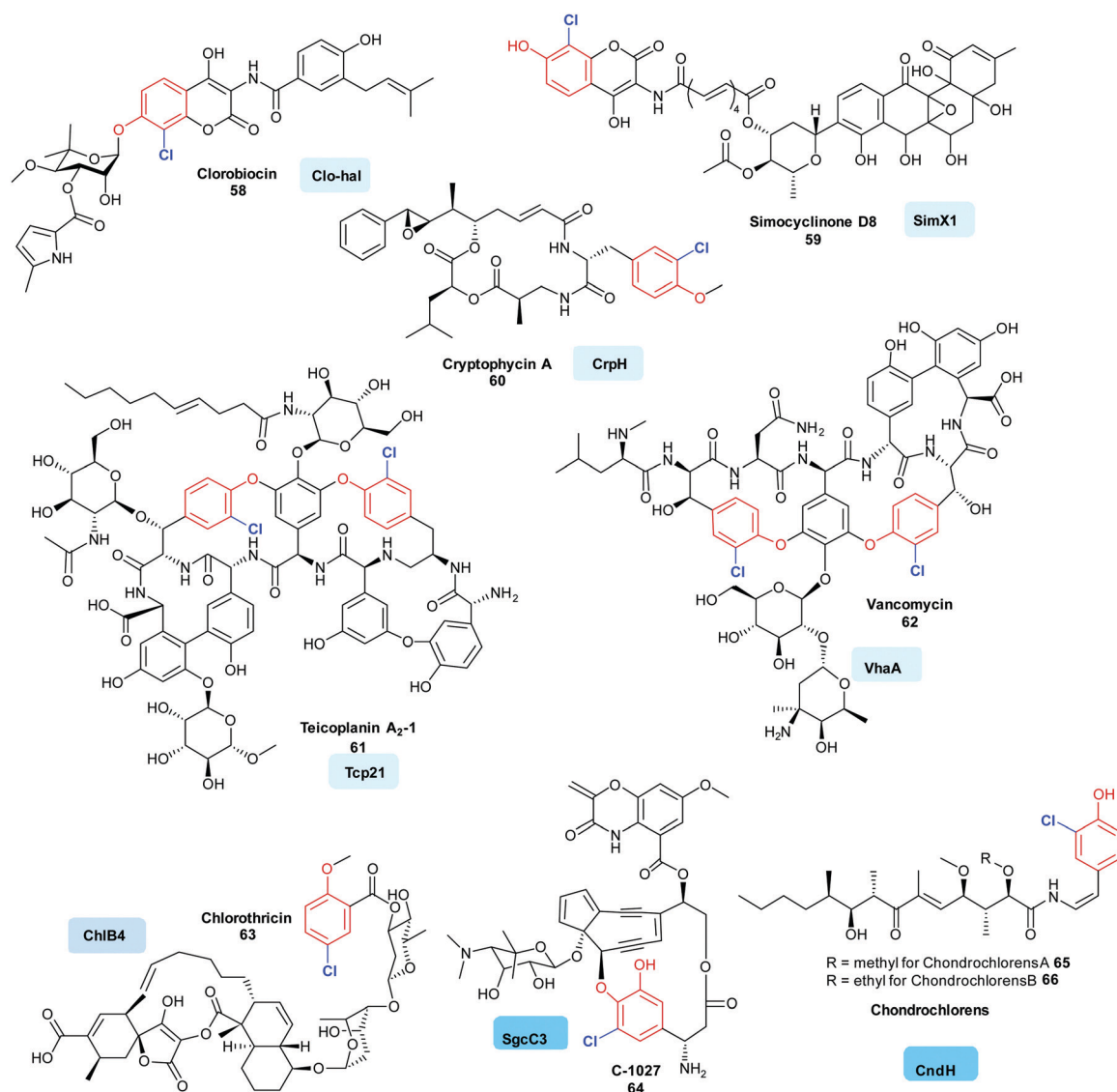


Scheme 5 Halogenation in the biosynthesis of the phenolic napyradiomycin **76** and merochlorins **77**, **78**: these natural products are synthesised from polyketide derived tetrahydroxynaphthalene (THN), the biosynthesis of these suites of metabolites involve versatile VHPOs and employ a chemical logic of chloronium formation promoting enantioselective intramolecular cyclisation.<sup>153,155</sup> Chlorination is indicated in blue.



respectively).<sup>133,134</sup> Evidence points to a similar halogenation strategy also being utilised in the biosynthesis of the peptidic cyanobacterial products aeruginosin (AerJ), cyanopeptolin (McnD), and cryptophycin A (CrpH) **60** of mixed PKS NRPS origin.<sup>135,140</sup> A similar biosynthetic logic is also predicted for the glycopeptide antibiotics (GPAs). Vancomycin **62** (VhaA) and teicoplanin **61** (Tcp21), are two potent GPAs; again, these compounds contain chlorinated tyrosine residues. In the case of the GPA teicoplanin **61**, careful *in vitro* investigations have shown that Tcp21 accepts the peptidyl carrier protein (PCP) supported aminoacyl substrate (tethered tyrosine), but will not accept the dipeptide or extended peptides as substrate.<sup>136</sup> In teicoplanin

biosynthesis, this amino acid residue is likely to be  $\beta$  hydroxylated at a later stage by a non-heme iron oxygenase, whilst in balhimycin biosynthesis OH-Tyr, and not tyrosine, is incorporated by the NRPS machinery.<sup>141</sup> Analogues of tyrosine may also be accepted by halogenases. SgcC3 accepts  $\beta$ -tyrosine, probably also in carrier protein tethered form, in the biosynthesis of the enediyne C-1027 **64**.<sup>137</sup> In a similar manner, enzyme tethered 3-hydroxy-4-methoxy-L-phenylalanine is seen to be processed to 2-chloro-3-hydroxy-4-methoxy-L-phenylalanine by Pep1 in the first steps of the biosynthesis of pepticinnamin G.<sup>142</sup> In chondrochloren **65**, **66** biosynthesis, CndH is indicated to process a different, perhaps decarboxylated yet enzyme tethered analogue of tyrosine.<sup>36</sup>



**Fig. 6** Representative bioactive natural products containing halophenol motifs, where the halogen is installed by a variant B FDH. Clorobiocin **58**,<sup>133</sup> simocyclinone D8 **59**,<sup>134</sup> cryptophycin A **60**,<sup>135</sup> are all postulated to be generated from the halogenation of an enzyme tethered tyrosine residue, in the first step (halogen added *ortho* in EAS fashion as mediated by Clo-hal, SimX1 and CrpH respectively). *In vitro* evidence supporting the early stage halogenation of an enzyme tethered tyrosine exists for teicoplanin **61** and vancomycin **62** (mediated by Tcp21<sup>136</sup> and VhaA).<sup>136</sup> C-1027 **64**<sup>137</sup> and the chondrochlorens **65**, **66** are both postulated to arise from the halogenation of an enzyme tethered analogue of tyrosine, in the first stages of the biosynthesis,<sup>36,138</sup> whereas a first step of chlorothricin **63** assembly is thought to be the halogenation of enzyme tethered salicylate (catalysed by ChlB4).<sup>139</sup> Chlorination is indicated in blue. Different regio-chemistries for chlorination are represented by different shades of blue.



Whilst many of the halogenases identified so far are seen to accept electron-rich aromatic amino acids in the biosynthesis of chlorothricin **63**, ChlB4 is postulated to act on the small polyketide subunit, 6-methyl salicylic acid<sup>139</sup> (Fig. 6).

**2.2.2 Late stage halogenation by variant A FDHs.** Whilst the variant B FDHs are not easily amenable to *in vitro* investigations, and are indicated to have a very limited substrate scope, the variant A FDHs are less limited by these problems and offer great opportunity for harnessing as biocatalysts. Amongst the most exciting of these, are enzymes that are reported to act at a late stage of the biosynthesis of structurally complex and sterically bulky phenolic/quinolic natural products. Rdc2<sup>143</sup> and RadH<sup>144</sup> share 87% sequence similarity; these two enzymes are identified to act in the final steps of biosynthesis of the fungal polyketide natural product radicicol **67**, an Hsp90 inhibitor, within the fungi *Pochonia chlamydosporia* and *Floropilus chiversii* respectively. Gsfl operates in the final stages of the assembly of griseofulvin **68**, a clinically utilised antifungal agent (Fig. 7).<sup>145</sup> TiaM from tiacumicin B biosynthesis operates in the final step of the biosynthetic pathway, functionalising a complex and glycosylated polyketide generating both the mono and dihalogenated products.<sup>226</sup>

CazI, from chaetoviridin A **68** biosynthesis, is postulated to operate on a complex phenolic intermediate, in the final stages of this cytotoxic polyketide.<sup>146</sup> In a similar manner, CtcP is seen to halogenate an advanced type II, aromatic polyketides core, in the biosynthesis of chlortetracycline **73**.<sup>147</sup> Heterologous pathway reconstruction studies indicate that in geodin biosynthesis GedL dichlorinates the advanced PKS metabolite sulochrin,

directly affording dihydrogeodin **72**.<sup>148</sup> Aspirochlorine **70** is an epidithiodiketepiperazine (ETP) toxin, produced by *Aspergillus orizae*, a mold used for millennia in Asian cuisine. The first stages of the biosynthesis of aspirochlorine involve the incorporation of two phenylalanines into a symmetric diketopiperazine (DKP). Intriguingly, one phenylalanine undergoes a C–C cleavage to convert it into glycine. The other phenylalanine is hydroxylated. AclH, operates in the final step of the biosynthesis, chlorinating the hydroxylated phenylalanine residue and affording the toxin.<sup>149</sup> XanH, a bifunctional FDH fused to its concomitant flavin reductase, has been shown capable of regioselectively chlorinating a complex late stage xanthone intermediate on the path to xantholipin **74**.<sup>150</sup> Pentachloropseudilin **71** has a series of interesting biological activities including inhibiting TGF-beta signalling, and impairing angiogenesis; yet its exact biosynthesis remains unclear. HalB was found in a cosmid library of the producing organism; the first halogenase identified from an actinomycetes. Notably, HalB shows 55% sequence identity to the pyrrole FDH PrnC. HalB has been shown to be capable, *in vitro*, of installing a single halogen into 2-(3,5-dibromophenyl pyrrole), but, whether or not it acts iteratively to install all 5 halogens in both the phenol and pyrrole ring in the natural system remains unclear.<sup>151</sup>

**2.2.3 Late stage halogenation by haloperoxidases.** Vanadium-dependent haloperoxidases (V-HPOs) have been identified that are involved in the biogenesis of halophenol containing natural products. The napyradiomycins (including napradiomycin B1 **76**) are polyketide-terpenoid natural products identified in marine streptomycetes, over 50 members of this suite of natural

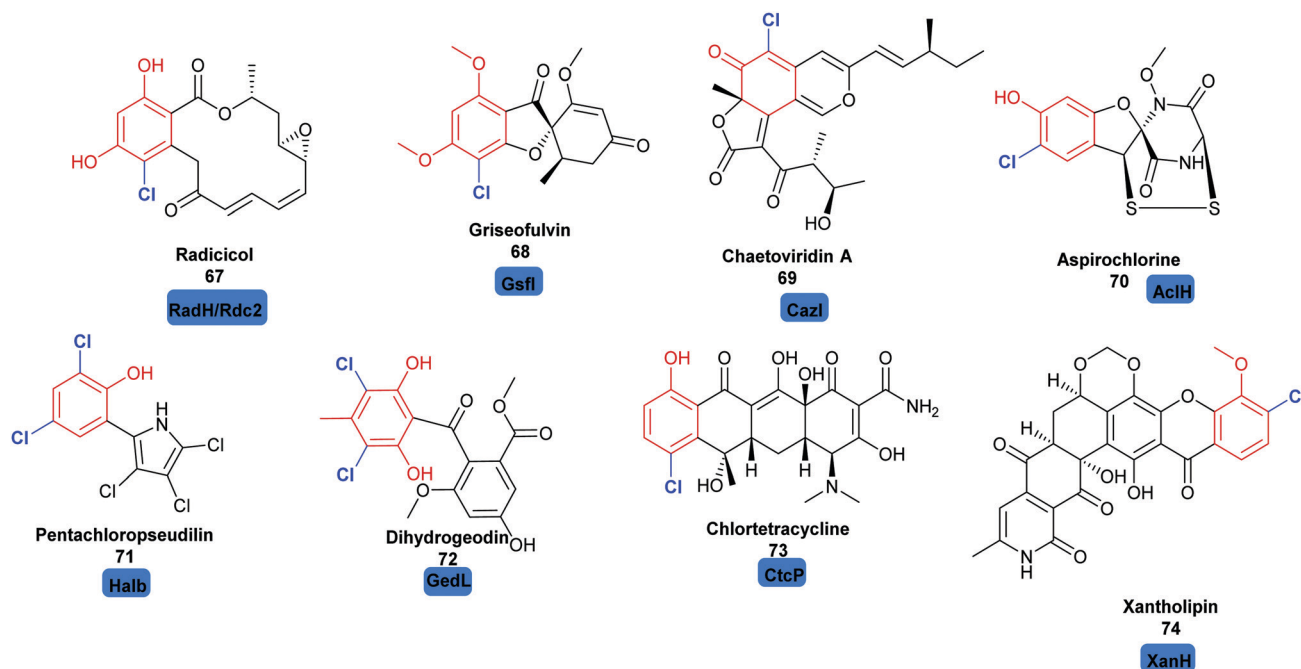


Fig. 7 Halogenated natural products where installation of halogen on a structurally complex non-tethered freestanding phenol is performed selectively by a halogenase: including radicicol **67**,<sup>143</sup> griseofulvin **68**,<sup>145</sup> chaetoviridin A **69**,<sup>146</sup> aspirochlorine **70**,<sup>149</sup> pentachloropseudilin **71**,<sup>151</sup> dihydrogeodin **72**,<sup>148</sup> chlortetracycline **73**,<sup>147</sup> and xantholipin **74**.<sup>150</sup> Chlorination is indicated in blue.



products have been identified so far. V-HPOs identified within the biosynthetic gene cluster include NapH1, and NapH4. NapH1 is particularly notable, this versatile enzyme acts upon geranylated tetrahydroxynaphthalene (THN) derivative, stereoselectively hydroxylating it whilst mediating regioselective chlorination (Scheme 5). NapT8 and NapH3 operate successively to catalyse prenylation and then an  $\alpha$ -hydroxy ketone rearrangement. NapH1 then operates once more to halogenate the newly introduced prenyl unit. The reaction is most likely to proceed *via* a halonium species, which is then opened by the hydroxyl that NapH1 previously introduced, thereby enabling enantioselective cyclisation. In a similar manner, NapH4 mediates halogenation of the geranyl unit, again most likely proceeding *via* the halonium species, and promoting intramolecular cyclisation.<sup>152</sup> The complex cyclic scaffolds of the merochlorins **77**, **78** are formed using similar chemical logic. Enzymatic total syntheses of napyradiomycin B1 **76** has been enabled in 18% yield.<sup>153</sup>

**2.2.4 Substrate scope: phenols.** Smaller, freestanding, non-native phenolic substrates have been demonstrated to be processed by variant A FDHs and HPOs (Fig. 8), revealing these enzymes to be useful biocatalysts. The employment of the HPOs, in this context, whilst enabling halogenation of sterically bulky substrates, results in mixtures of regioisomers and poly-halogenated compounds. This is a direct result of most HPOs releasing free hypohalous acid; in this context, the enzyme has little control over the regiochemistry of the transformation, and it is simply the most electron-rich region(s) of the substrate that are halogenated. In contrast, the FDHs usually accept much smaller substrates but are generally highly regioselective. PltM has been shown capable of mono **117** or di-halogenation **118** of phloroglucinol (Fig. 8J), as well as being able to accept resveratrol as a substrate.<sup>51</sup> Wild type RebH, Gsfl, and Thal have been shown capable of processing a broad range of non-natural substrates containing phenolic moieties.<sup>154</sup> Exceptions may be seen with RadH/Rdc2 and VirX1, FDHs with much more open and accessible active sites, capable of regioselectively monohalogenating much larger phenolic substrates.<sup>47,143,144</sup>

**2.3 Pyrrolic substrates.** Substituted pyrroles are also widely used in medicine and agrochemistry. In nature, halogenases may be found that are capable of mono- or poly-halogenating pyrroles. Halopyrrole containing natural products include pyrrolnitrin **24**,<sup>34</sup> pentachloropseudilin **71**,<sup>151</sup> chlorizidine A **131**,<sup>162</sup> hormaomycin **132**,<sup>163</sup> pyoluteorin **133**<sup>164</sup> and pyrrolomycins **134–137**<sup>165</sup> (Scheme 6). PrnC, a variant A FDH, chlorinates the pyrrole of **128** in the final step of pyrrolnitrin **24** biosynthesis.<sup>34</sup> Here, the pyrrole is generated through the cleavage and rearrangement of 7-chlorotryptophan **127** (Scheme 6). HalB, which shares 45% sequence identity with PrnC, is implicated in halogenating the pyrrole of pentachloropseudalin **71**.<sup>151</sup> For chlorizidine a **131**, hormaomycin **132**, marinopyrrole **138**,<sup>166</sup> the pyrrole is formed by oxidation of a proline uploaded onto an acyl carrier protein (ACP), the ACP tethered pyrrole is then halogenated. Bmp2 is one such variant B flavin-dependent pyrrole halogenase, mediating the

intriguing tetrabromination of the PCP-bound pyrrole ring, with the introduction of the fourth bromide, to C1, apparently triggering thioesterase mediated release and decarboxylation to restore aromaticity.<sup>48</sup> Curiously, the biosynthetically-related halogenase, Bmp5, employs a chemically similar strategy of decarboxylative bromination in the conversion of *p*-hydroxybenzoic acid to 2,4-dibromophenol (see Section 1.1.2). Mpy16 from marinopyrrole **138** biosynthesis is structurally similar to Bmp2, yet stops at dihalogenation,<sup>166</sup> as do Clz5<sup>162</sup> and PltA<sup>164</sup> in chlorazidine A **131** and pyoluteorin **133** biosynthesis, respectively. Pyr29 is also postulated to di-halogenate an ACP tethered pyrrole **130** in the biosynthesis of the pyrrolomycins,<sup>165</sup> whilst HrmQ from hormaomycin **132** biosynthesis mediates monohalogenation.<sup>163</sup> Structure-guided mutagenesis has been performed on wild-type Bmp2 to generate a Bmp2 triple mutant (Bmp2-TM) which proceeds no further than dibromination.<sup>48</sup>

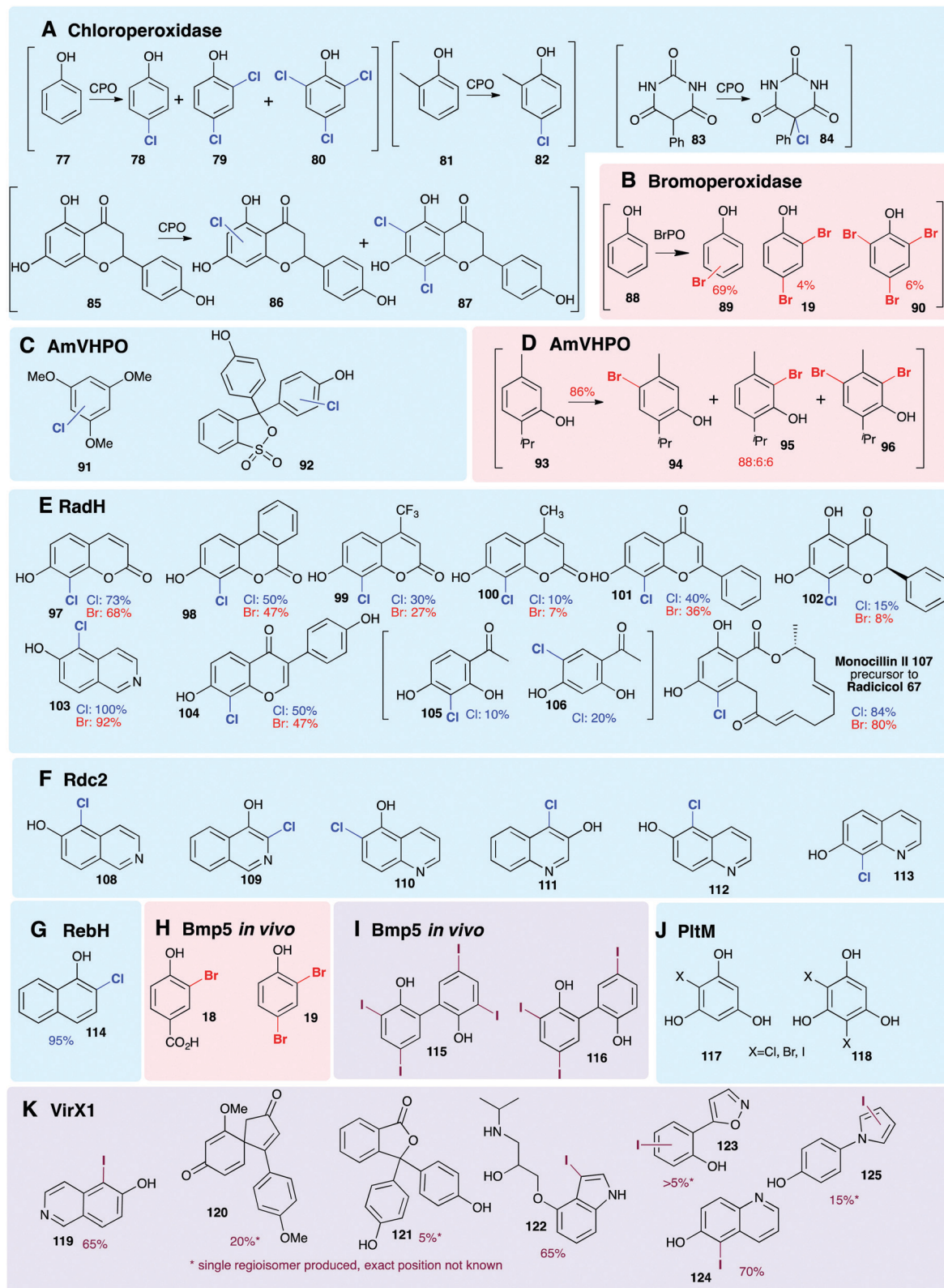
**2.4 Aromatic substrate scope beyond indole, phenol and pyrrole.** Though the vast majority of natural substrates for the FDHs and haloperoxidases studied so far are indolic, phenolic or pyrrolic in nature, they have been effectively employed in installing halogens into other substrate classes. Generally, once again, the FDHs are seen to confer greater regio- and substrate specificity than the HPOs that generate and release free HOX. Initial reports of wild-type flavin-dependent halogenases had been indicative of a fairly modest and narrow substrate scope. However, RebH, PrnA, PyrH, and SttH have been shown capable of processing a series of anilines including kynureide **171**, anthranilamide to give **172** and, to a lesser extent, anthranilic acid to give **174** (Fig. 9C–F).<sup>79,80</sup> These studies also gave insight into factors governing regiospecificity. With PyrH, regiospecific *para*-chlorination **171** of kynurenine was observed whereas PrnA mediated the formation of the *ortho*-chlorinated product **166** exclusively, a product that is more chemically challenging to access. PyrH could also be used to achieve the regiospecific *para*-chlorination **172** of anthranilamide, whereas use of PrnA resulted in a 86% : 14% *para* : *ortho* mixture. Mutations could be effectively employed to modulate *ortho*:*para* selectivities and improve yields for these unnatural substrates.<sup>79</sup> Chlorination of naphthalen-2-amine to give **165** by RebH *ortho* to –NH<sub>2</sub> is achieved in a high yield of 93%.<sup>126</sup>

The halogenation of sterically bulky substrates can be enabled using haloperoxidases. As there is generally no substrate binding site, and free HOX is released, the substrates that can be halogenated using these systems are not limited by size. However, as is usually expected for an HPO that releases HOX, series of regioisomers, as well as mono, di, and tri halogenated compounds result (Fig. 9G).<sup>169</sup> For example, heme iron-dependent chloroperoxidase (CPO) from *C. fumago* has been used to give chlorinated analogues of naphthalene **182**, anthracene **183** and pyrenes **185** and **186**. A vanadium-dependent bromoperoxidase from the marine algae *Ascophyllum nodosum* has also been employed mono and dibromination 4,6,8-trimethylazulene **179** (Fig. 9G), though the instability of the resultant product hindered purification.<sup>168</sup>

To achieve regioselective electrophilic halogenation, generally FDHs are needed, however, historically these have shown fairly



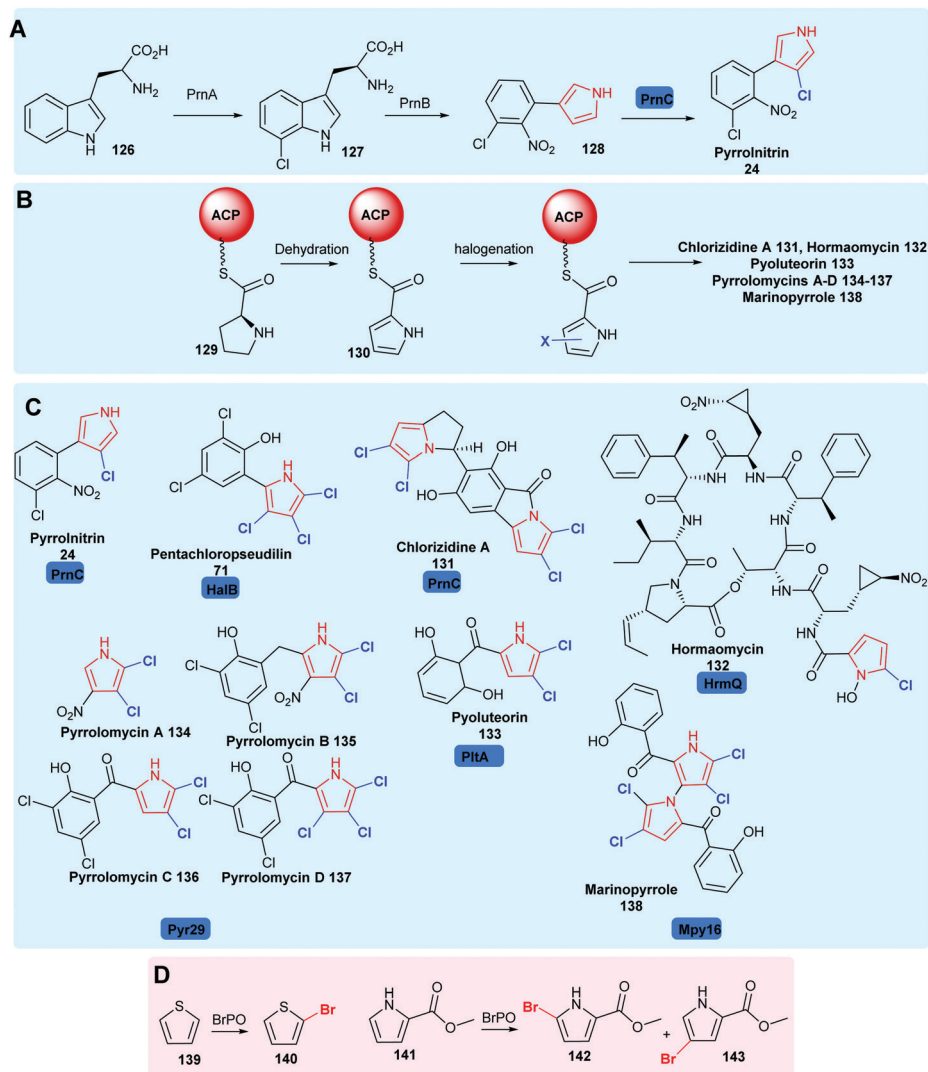




**Fig. 8** Representation of some of the diverse phenols that are accepted as non-natural substrates by different wild-type halogenases. All experiments are *in vitro*, unless specified otherwise, and % conversions reported as determined by LC/LC-MS. Enzymes include (A) chloroperoxidase (heme iron-dependent),<sup>156–158</sup> (B) bromoperoxidase (vanadium-dependent)<sup>159</sup> (C and D) AmVHPO (vanadium-dependent)<sup>107</sup> (E) RadH (FDH) chlorination,<sup>144</sup> (F) Rdc2 (FDH) chlorination and bromination,<sup>143,160</sup> (G) RebH (FDH) chlorination,<sup>126</sup> (H and I) Bmp5 brominase (FDH) *in vivo*,<sup>45</sup> (J) PltM (FDH) associated with transcriptional regulation of pyoluteorin **43** biosynthesis,<sup>51,161</sup> (K) VirX1 (FDH from a virus, the first FDH with a preference for iodination).<sup>47</sup> Chlorination, bromination and iodination are colour coded blue, red and purple respectively.







**Scheme 6** Naturally occurring halopyrrole containing compounds. (A) The biogenesis of pyrrolnitrin **24** involving a variant A FDH PrnC. (B) The enzyme tethered biosynthesis of halopyrroles from proline, employing variant B FDHs. (C) Bioactive compounds containing halopyrroles including pyrrolnitrin **24**,<sup>34</sup> pentachloropseudilin **71**,<sup>151</sup> chlorizidine A **131**,<sup>162</sup> hormamycin **132**,<sup>163</sup> pyoluteorin **133**,<sup>164</sup> pyrrolomycins **134–137**<sup>165</sup> and marinopyrrole **138**,<sup>166</sup> noting the halogenases utilised in their construction. (D) Biocatalytically vanadium-dependent bromoperoxidases may be utilised to monobrominate thiophene and pyrroles.<sup>167</sup> Chlorination is colour coded blue.

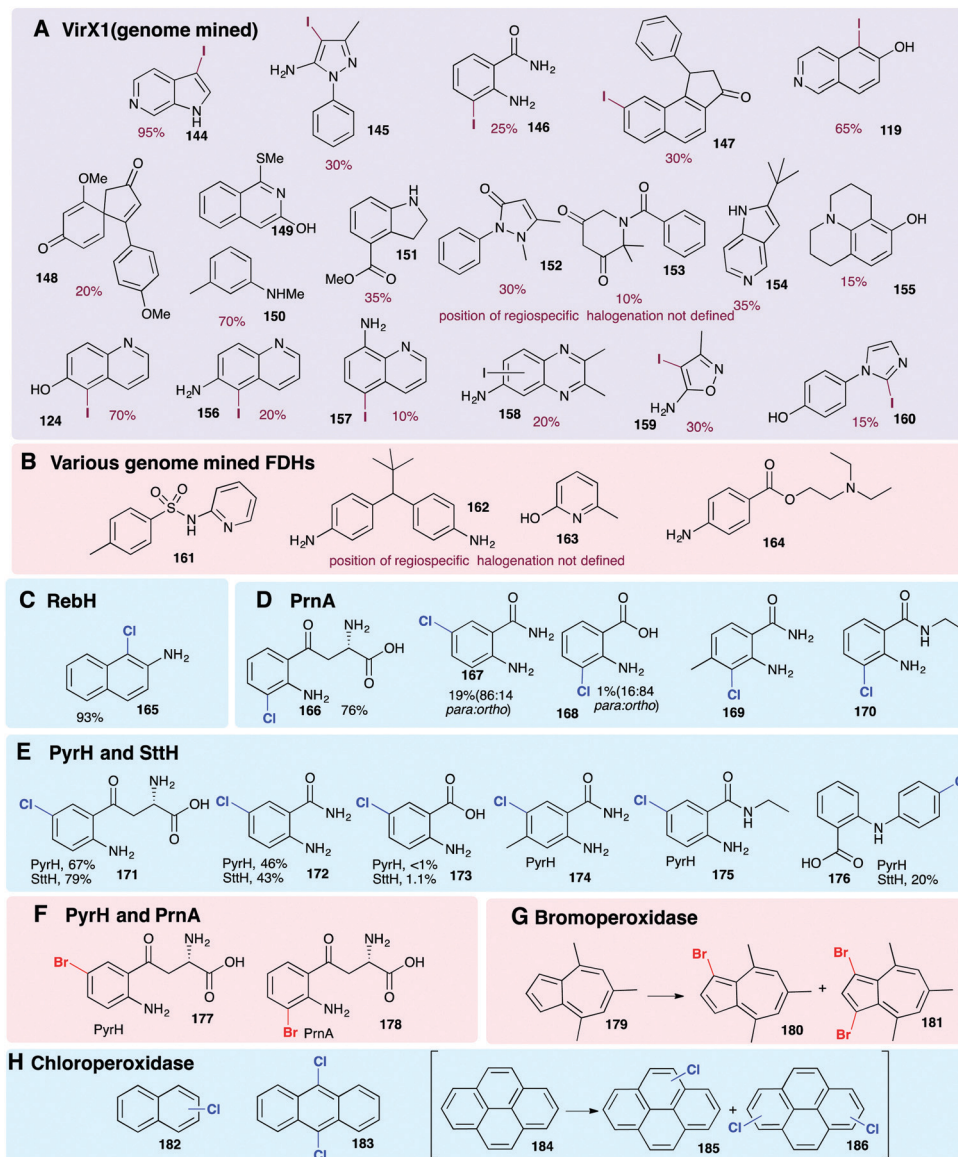
narrow substrate scope. In 2019, new FDHs were revealed that had been found using *in silico* genome mining approaches (Fig. 9A and B). These halogenases were able to accept much larger substrates and showed a much broader substrate scope than any FDHs examined previously. VirX1, with a preference for iodination, was demonstrated to be capable of regiospecifically iodinating or brominating a diverse portfolio of substrates with a variety of steric and electronic demands, and some of which might be considered to be only weakly activated. Good conversions of a wide range of unnatural substrates are possible with this unusual enzyme from a virus.<sup>47</sup> Similar informatics-led approaches have been utilised by the Sewald and Lewis groups, further demonstrating that natural FDHs do exist with broader substrate scope.<sup>170</sup>

**2.5 Aliphatic substrate scope for biocatalytic halogenation.** The NHFeHals functionalise unactivated  $\text{sp}^3$  carbons. The

selective C–H activation that these enzymes are capable of is challenging to achieve using even modern synthetic methodologies. Prior to 1998, these radical halogenases were unknown, however, the observation that the cyanobacterial natural product, barbamide **189**, contained a leucine residue with one trihalogenated methyl group, led Willis and co-authors to postulate that the enzymes involved in its biosynthesis perhaps proceeded *via* a radical mechanism.<sup>173</sup> Over half a decade later, SyrB2, from the biosynthesis of syringomycin E **182**<sup>46</sup> by *Pseudomonas syringae* B301D, was the first member of this new family of halogenases to be biochemically and structurally characterised (Scheme 7).

Like the FDHs, the NHFeHals may be subdivided into those that operate only on protein tethered substrates (which we call here the variant B NHFeHals, to be consistent with the variant B FDHs that also operate only on protein tethered substrates), and the perhaps more biotechnologically useful, and recently





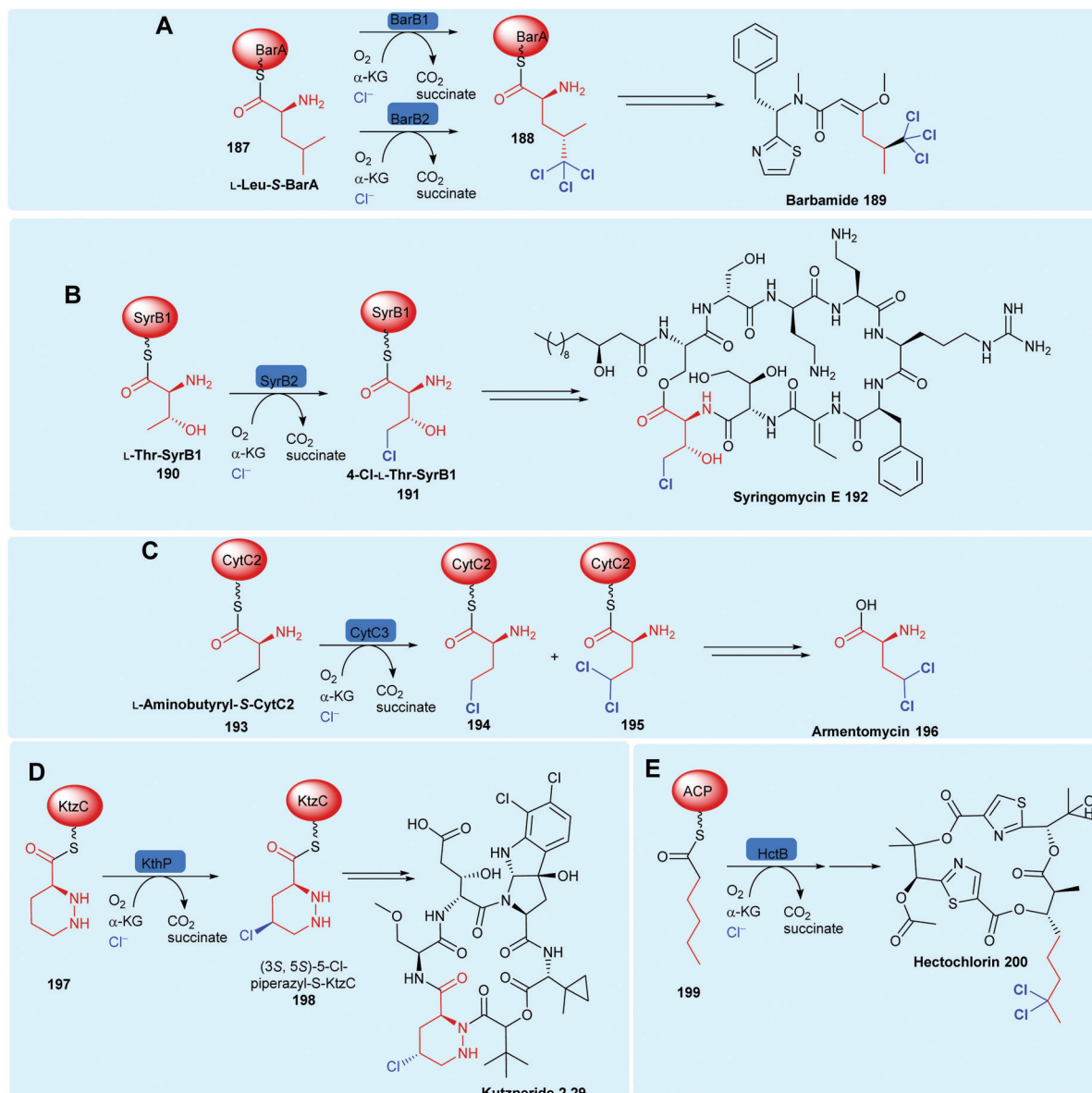
**Fig. 9** Representative selection of aromatic moieties (beyond indole, pyrrole and phenol) that may be accepted as substrates for enzymatic halogenation. Regiochemistry and conversions are given where reported in the primary literature. Enzymes include (A) genome mined VirX1, a variant A FDH with a very broad substrate scope, the first halogenase to be isolated from a virus and the first FDH to show preference for iodination.<sup>47</sup> Notably, the regioselective halogenation of a diverse range of both small and large substrates may be seen including several less electron-rich and less activated substrates (bromination of all substrates also possible). (B) A series of genome mined FDHs, again, halogenation of sterically more bulky substrates may be seen (bromination of all substrates also possible).<sup>154</sup> (C–F) Anilines and anthranilates processed by variant A FDHs RebH, PrnA, PyrH and SttH, respectively.<sup>79,80,126</sup> (G) Vanadium-dependent bromoperoxidase capable of brominating bulky substrates.<sup>168</sup> (H) Heme-iron dependent chloroperoxidases shown to be capable of processing bulky, planar modestly activated compounds such as pyrenes, mixtures of regiochemistries and levels of substitution are observed.<sup>169</sup> Chlorination, bromination and iodination are colour coded blue, red and purple respectively.

discovered variant A NHFeHals that process free substrates (see Section 1.2). The variant B type NHFeHals have proved challenging to handle requiring anaerobic conditions, nevertheless they catalyse series of intriguing chemistries, including cryptic halogenation *en route* to the installation of cyclopropyl or alkynyl motifs, notable examples may be seen in jamaicamide **204**,<sup>61</sup> curacin A **208**,<sup>174</sup> coronatine **213**<sup>175</sup> and kutzneride **29** biosynthesis (Scheme 8).<sup>118</sup> The need for covalent substrate tethering prior to halogenation means that variant B

halogenases have less biotechnological potential at the current time, and will be discussed just briefly in the following section.

**2.5.1 Tethered substrates with unactivated  $sp^3$  carbon centres.** NHFeHals capable of processing a variety of very different tethered substrates are known and have been shown to be capable of selectively functionalising terminal methyl groups and internal methylenes. Tri-chlorination of the pro-*R* methyl of peptidyl carrier protein (BarA) tethered leucine **187** by



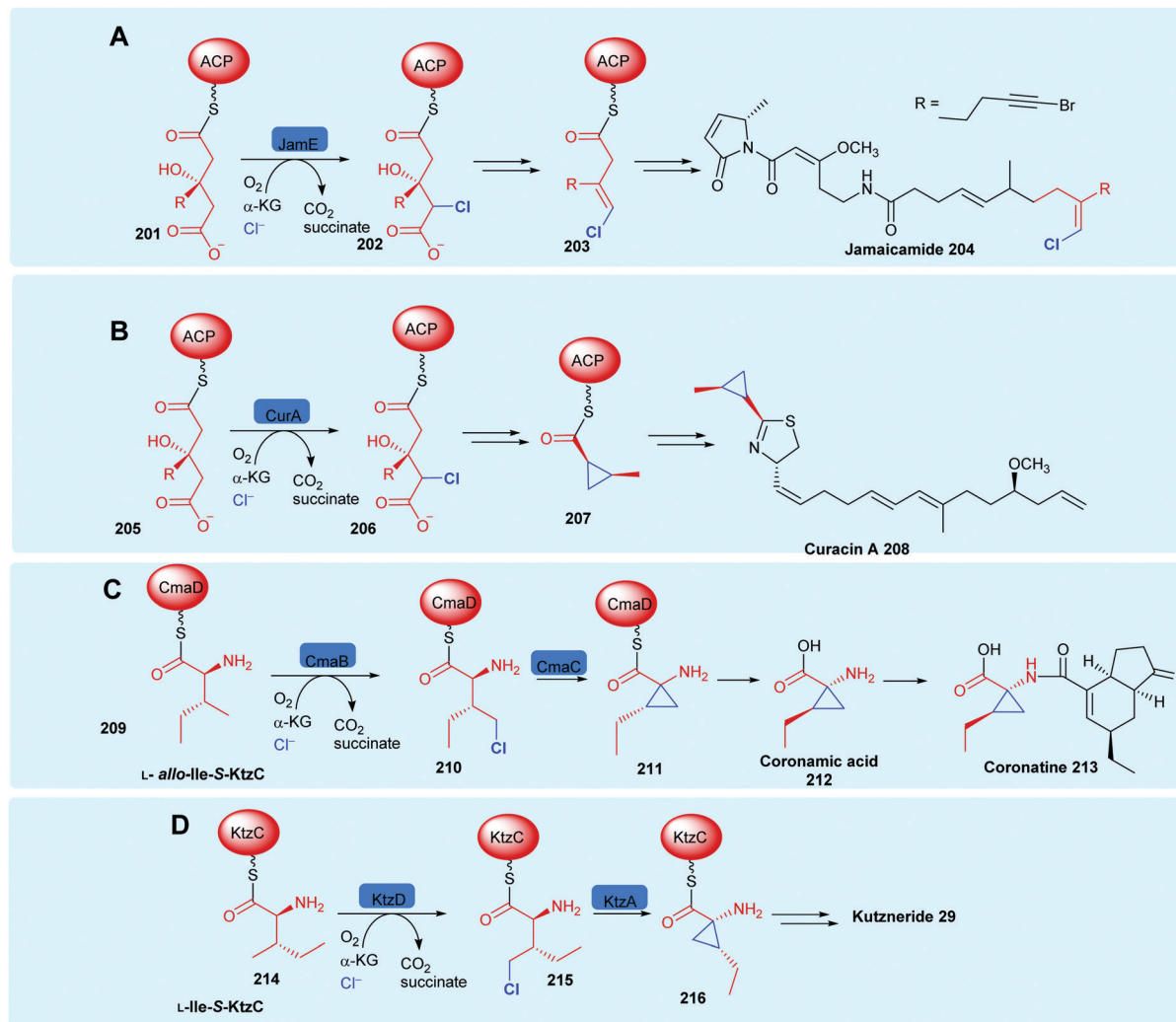


**Scheme 7** Bioactive compounds where installation of halogen on a protein-tethered aliphatic substrate is performed selectively by a NHFeHal, including (A) barbamide **189**,<sup>60</sup> (B) syringomycin E **192**,<sup>58</sup> (C) armentomycin **196**,<sup>171</sup> (D) kutzneride 2 **29**<sup>118</sup> and (E) hectochlorin **200**.<sup>172</sup>

tandem action of BarB1 and BarB2,<sup>60</sup> and di-chlorination and mono-chlorination of the  $\gamma$ -methyl of peptidyl carrier protein tethered L-2-aminobutyric acid **193** and L-threonine **190** by CytC3 and SyrB2 has been demonstrated through *in vitro* reconstitution experiments by Walsh and co-workers (Scheme 7).<sup>58,171</sup> CytC3 and SyrB2, have a level of similarity (58% identity, 71%) similarity and process these subtly different substrates. Beyond amino acids, NHFeHals are known that can process tethered piperazines, fatty acids and tethered polyketide intermediates. Examples include, KthP, which mediates the generation of the 5-chloropiperazyl motif in kutzneride **29** biosynthesis,<sup>118</sup> and HctB responsible for chlorination at C5, of an acyl-carrier protein tethered hexanoate **199**, in the biosynthesis of the antifungal hectochlorin by cyanobacterium *Lyngbya majuscula* (Scheme 7).<sup>172</sup>

**2.5.2 Unactivated  $sp^3$  carbon centres, processed as free, non-tethered substrates: enzymatic generation of alkyl halides.** The discovery of WelO5 from *Hapalosiphon welwitschii* UTEX B1830, the welwintindolinone **A 219** producer, by Liu and coworkers was ground-breaking; for the first time, a NHFeHal could be seen to process non-enzyme tethered substrates. Its regioselective and stereoselective halogenation of a complex carbocycle, renders it an exciting tool for biotechnological applications. WelO5 monochlorinates 12-epi-fischerindole **U 217** and 12-epi-hapalindole **C 220** (Scheme 12),<sup>62</sup> and has been demonstrated to enable bromination of the same substrates, leading to their enhanced antibacterial activity.<sup>176</sup> Sharing 79% sequence identity with WelO5, AmbO5 was subsequently discovered through the analysis of the ambiguine **53** biosynthetic gene cluster.<sup>83,177</sup> AmbO5 accepts a wider scope of substrates than WelO5 and





**Scheme 8** Installation of halogen on protein-tethered aliphatic substrate performed selectively by a NHFeHals, in the biosynthesis of natural products, enabling alkene generation and cyclopropane generation. (A) Jamaicamide **204**,<sup>61</sup> (B) curacin A **208**,<sup>174</sup> (C) coronatine **213**,<sup>175</sup> and (D) kutzneride **29**.<sup>118</sup> Notable similarities may be observed between the first steps in jamaicamide, vinylchloride formation and curacin cyclopropyl biosynthesis. Coronamic acid and kutzneride **29** biosynthesis are initiated by the halogenation of the  $\gamma$ -methyl of different diastereoisomers of L-isoleucine.

selectively chlorinates a range of fischerindole (**52**, **53**), hapalindole (**222**) and ambiguine alkaloids (**56**) (Scheme 9A).<sup>83</sup> By generating hybrid enzymes consisting of the N-terminus of WelO5 and C-terminus of AmbO5 (and vice-versa), some success of halogenating a slightly wider scope of hapalindole type molecules was achieved.<sup>83</sup>

Initial sequence and mutational analysis of WelO5 and AmbO5 pointed to eleven C-terminal residues, likely influencing substrate utilisation.<sup>180</sup> The potential to develop WelO5 through directed evolution to improve the biocatalytic capabilities and extend substrate scope has been attempted by a number of groups. Hoebenreich and coworkers used structure-guided direct evolution to develop variants of WelO5 homologue Wi-WelO15 from *Westiella intricata* HT-29-1 capable of selectively chlorinating non-natural hapalindole **235**, and hapalindoles **237–240** containing a ketone moiety at the position of the natural isonitrile.<sup>178</sup> Buller and coworkers engineered variants of another homologue WelO5\* for selective halogenation at two

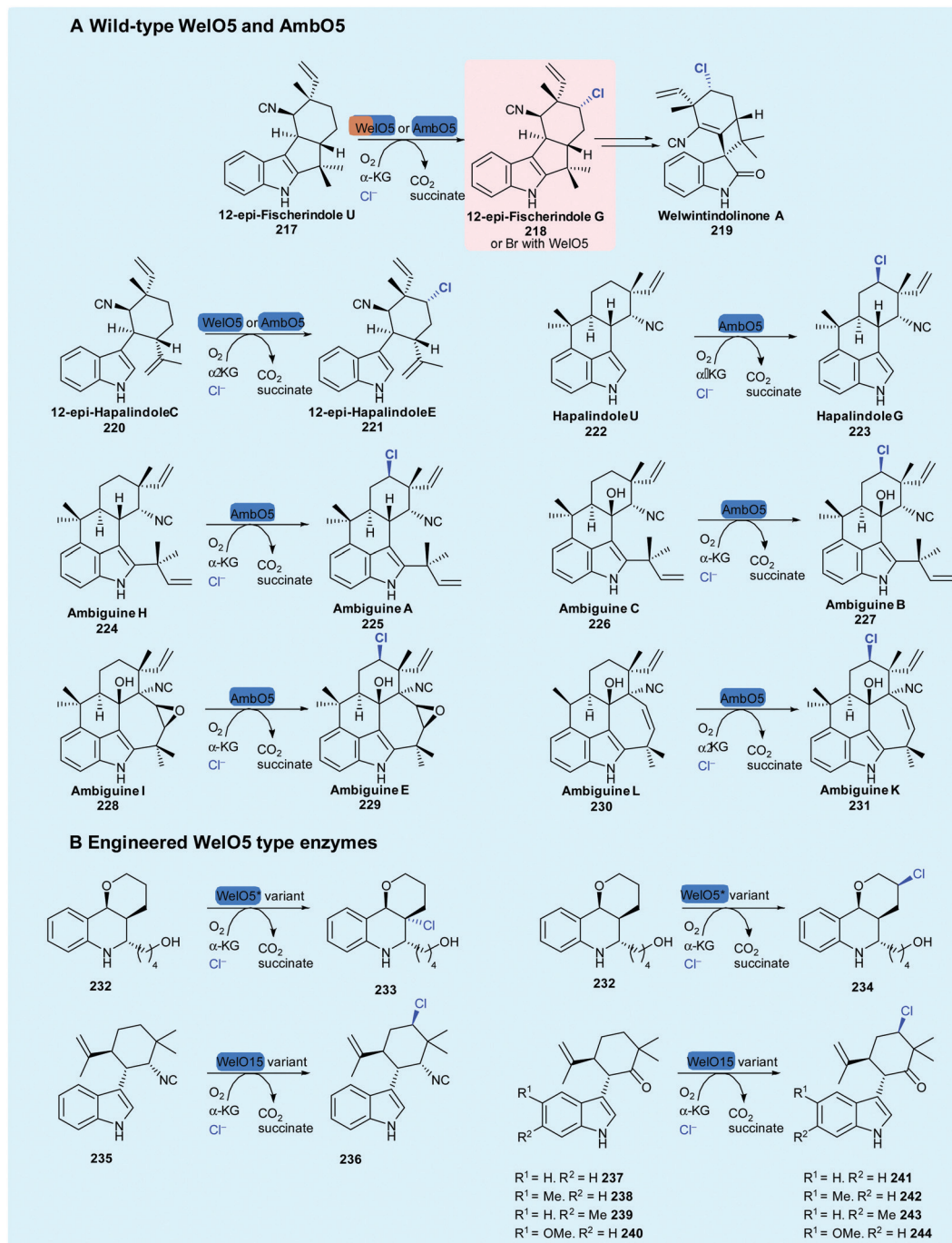
separate positions on a martinelline analogue **232**,<sup>179</sup> representing the first successful biocatalytic installation of a halogen by WelO5 type enzymes on a substrate that is notably different to their natural hapalindole targets. Both Buller and Hoebenreich identified amino acid positions, which seem to play an important role in the regioselectivity of these enzymes towards non-natural substrates, opening up possible opportunities for rational re-engineering of these catalysts.

The co-crystal structure of WelO5 and native substrate was used in a structural search for other enzymes capable of alkyl halogenation of different substrates. In this manner, the hydroxylase SadA was identified. Minor modification to the active site and coordination of the FeII (D157GSadA) resulted in an enzyme capable of chlorinating as well as hydroxylating its natural substrate.<sup>182</sup>

A second series of variant A NHFHals that operate on substrates that are not covalently tethered to a carrier protein (free substrate NHFe halogenases) have recently been







Scheme 9 Known substrate scope of wild-type NHFeHals WelO5<sup>62</sup> and AmbO5<sup>83</sup> and engineered varieties of WelO15<sup>178</sup> and WelO5\*<sup>179</sup>

discovered, and named the BesD family. Like CurA, CmaB, KtzD (Scheme 8) and JamE, BesD is a cryptic halogenase, and was discovered in the biosynthetic pathway of the amino acid  $\beta$ -ethynylserine **247** (Bes) from *S. cattleya*, after knockouts of non-essential amino acid desaturases pointing towards a novel enzymatic production of the alkyne moiety.<sup>183</sup> Purified BesD was shown to selectively chlorinate free lysine **245**, producing 4-Cl lysine **246**.<sup>84</sup> Subsequent C–C cleavage and elimination of the installed chlorine was shown to lead to the terminal alkyne (Scheme 10A). BesD has low sequence identity to both substrate

bound SyrB2 (7%) and variant B prototype WelO5 (11%), instead having a much higher identity to predicted hydroxylases (>46%). Rather than being a standalone enzyme, BesD was shown to be part of a variant A NHFeHals.<sup>84</sup> Homology networks were produced from hits arising from a sequence-based homology search of BesD. This approach resulted in the discovery of 20 more halogenases that act on free amino acid substrates. The range of substrates was diverse, with BesD and other enzymes accepting lysine but also ornithine, while the hydrophobic amino acids leucine, isoleucine, and norleucine,

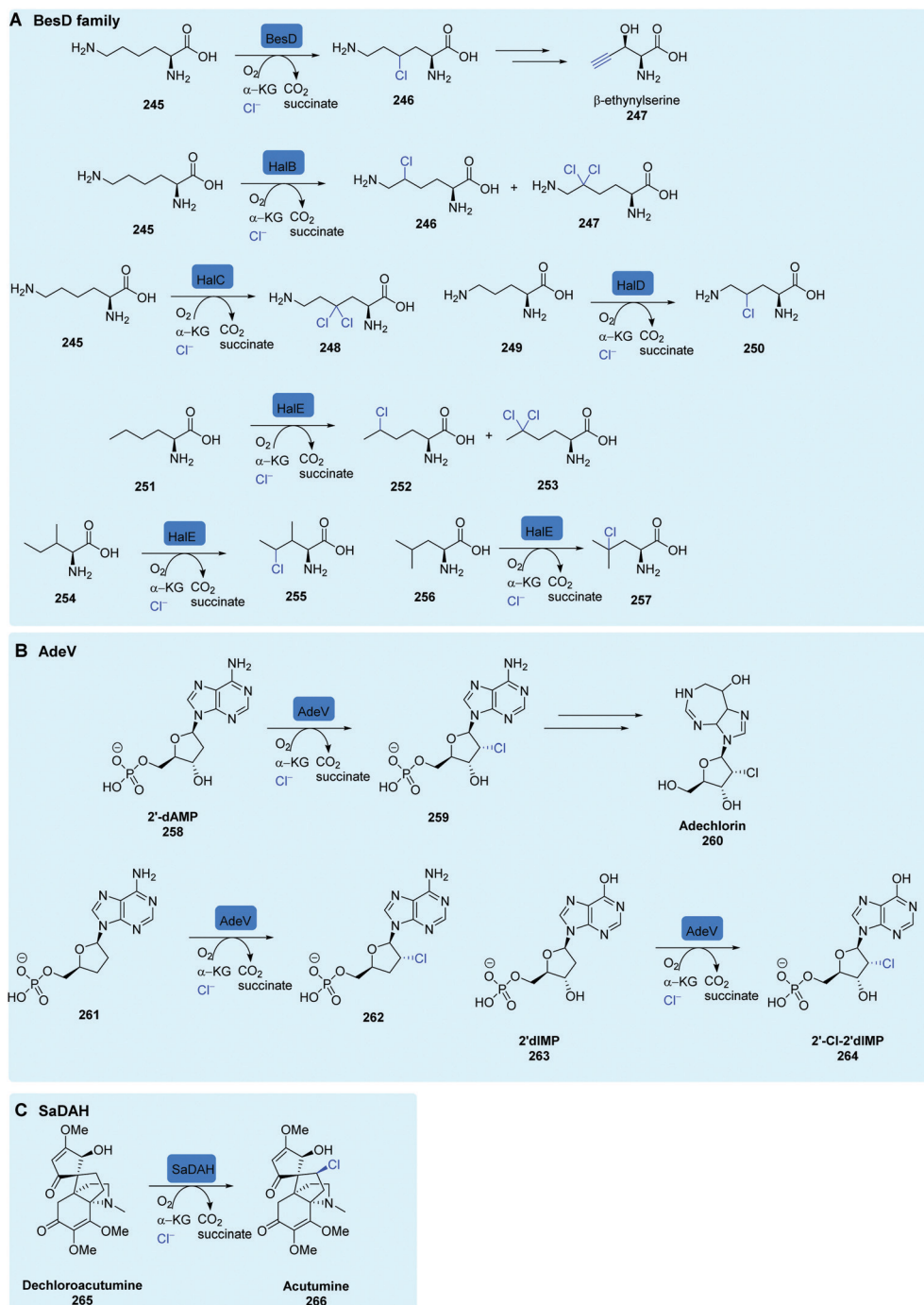




were shown to be substrates for PrHalE (Scheme 10A). A selection of these halogenases was shown to perform alongside amino acid metabolising enzymes to produce chlorinated heterocycles, diamines, and  $\alpha$ -keto acids from lysine (SwHalB), ornithine (PkHalD), and norleucine (PrHalE). Remarkably, all enzymes retained fidelity for halogenation over hydroxylation, while the radical halogenase HalB (different to the FDH HalB) from *Streptomyces wuyuanensis* (SwHalB) was able to accept

bromine and azide anions to produce bromo-lysine and azido-lysine. Chloro-lysine, generated by SwHalB and PfHalA, could also be incorporated into a 9 amino acid peptide using an *in vitro* transcription and translation system-suggesting a potential use for these enzymes in production of more complex natural product analogues.<sup>84</sup>

Recently, variant A NHFHal capable of halogenating nucleosides, named AdeV, has been discovered in the biosynthesis of



**Scheme 10** NHFe halogenases. (A) Known substrate scope of the NHFe halogenase BesD and enzymes from the BesD family,<sup>84</sup> (B) AdeV chlorination of 2'-dAMP in adechlorin biosynthesis and other known substrates of AdeV,<sup>85</sup> (C) SaDAH chlorination in acutumine biosynthesis.<sup>181</sup>

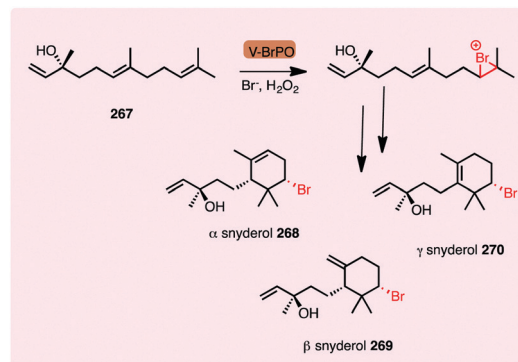


the chlorinated natural product adechlorin **260** in *Actinomadura* sp. ATCC 39365, opening up the path to access valuable halonucleosides<sup>85</sup> (Scheme 10B). AdeV was shown to have 15% similarity to WelO5. Gene knockout experiments confirmed its role in adechlorin biosynthesis. AdeV acts in the early stages of biosynthesis on free nucleoside 2'-deoxyadenosine-5'-monophosphate (2'-dAMP) **258** to generate 2'-Cl-2'-dAMP. *In vitro* assays revealed that 2'-deoxyadenosine (2'-dA), identical to 2'-dAMP apart from the 5'-phosphate moiety, was not accepted as a substrate, indicating that the presence of this phosphate is essential for substrate binding and halogenation activity. Consistent with this logic, two other phosphorylated nucleosides 2',3'-dideoxyadenosine monophosphate **261** and 2'-deoxyinosine-5'-monophosphate (2'-dIMP) **263**, were accepted by AdeV and converted to their chlorinated counterparts indicating a level of enzyme promiscuity. However, the natural substrate 2'-dAMP showed the highest levels of conversion.

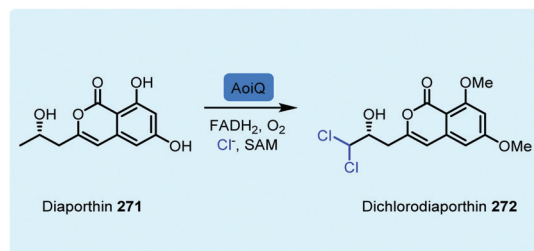
Several plant species are known to generate halogenated metabolites. The toxin fluoroacetate is produced in many plants including *Camellia sinensis* (from which tea is generated). Fatty acid derivatives of fluoroacetate such as  $\omega$ -fluorooleic acid accumulating in the seeds of *Dichapetalum toxicarium*.<sup>181</sup> An exciting new NHE halogenase DAH that performs the stereo-selective late stage chlorination of the complex alkaloid dechloroacutumine **265**, (Scheme 10C) produced by menispermaceae plants, has been found and its activity demonstrated *in vitro*. This represents the very first NHEHal found in plants.<sup>181</sup> Phylogenetic analysis indicates that DAH, which has a variant in *Menispermaceae canadense* (McDAH) and *Sinomenium acutum* (SaDAH), evolved independently from the previously discussed bacterial NHEHalS, potentially being an example of parallel evolution in halogenated natural product metabolism. *In vitro* assays against a wide range of alkaloids indicated that DAH was highly selective towards its natural substrate, with no other small molecules being accepted for halogenation. As with some other halogenases, DAH was shown to accept azide anions as well as halides, and convert (–)-dechloroacutumine to 11-azido-dechloroacutumine.

Beyond the NHE halogenases, alkyl halides can be generated by a series of vanadium-dependent haloperoxidases. An elegant cyclisation cascade is initiated by the selective bromination of a single alkene within a terpenoid; the resultant bromonium undergoes ring opening affording the alkyl bromide. For example, the vanadium bromoperoxidase from *C. officinalis* brominates the terpenoid precursor (*E*)-(+)-nerolidol **267**, yielding cyclised snyderols **268–230**.<sup>184</sup> For  $\beta$ - and  $\gamma$ -snyderol, a single diastereoisomer is produced. This provides an exciting biocatalytic opportunity as synthetic methods have succeeded only in forming a mixture of two diastereoisomers of each product (Scheme 11). In a similar manner, chlorination, mediated by haloperoxidases, may be seen to generate chloronium species, initiating cyclisation cascades. Installation of chlorine onto the aliphatic moiety of SF2415B1 by NapH1 allows cyclisation of the structure to afford a cyclic ether, offering potential for synthesis of napradiomycins **76** (Scheme 5) with cytotoxic and antibacterial properties.<sup>155</sup>

Notably, a small number of FDHs process aliphatic compounds. One such example is CmlS, an FDH from chloramphenicol



Scheme 11 V-BrHPO catalysed terpene bromination and cyclisation event in snyderol biosynthesis.<sup>184</sup>



Scheme 12 FDH AoiQ dichlorination and dimethylation of diaporthin **271**.<sup>186</sup>

biosynthesis implicated, through *in vivo* studies, in the generation of a dichloroacetyl moiety.<sup>185</sup> In this unusual FDH, the flavin cofactor is covalently linked to the enzyme. The malonyl-CoA substrate is proposed to form an enolate, stabilised through hydrogen bonding to tyrosine, this species may then react with a proximal chloramine.<sup>44</sup> Beyond the halogenation of tethered malonyl-CoA by CmlS, other exciting observations of the applications of these enzymes are noted. Intriguingly, an FDH from a fungus, AoiQ, has been demonstrated to regioselectively chlorinate an unactivated terminal  $sp^3$  carbon of a freestanding molecule **271**, in the biosynthesis of orthosporin (Scheme 12). So far, AoiQ is unique amongst other flavin-dependent halogenases in its ability to perform this task. Although the mechanism for this reaction is unknown, it has been postulated by Hertweck and coworkers that, though it is a FDH, it may proceed *via* a radical reaction.<sup>186</sup> Notably, like KrmI, it is one of the few fused halogenases to be explored. In the case of AoiQ, it contains a functional O-methylation domain. The mystery as to how it operates remains to be revealed.

### 3. Approaches toward improved biocatalytic halogenation, and steps toward upscaling

Considerable effort has been invested into engineering and evolving the biocatalytic activity of halogenases. Studies have predominantly focussed upon the FDHs in order to alter substrate profiles, modify regiochemistries of halogenation and enhance biocatalytic stability. Greater structural understanding of these enzymes has enabled progress to be made in their rational redesign.<sup>187,188</sup> Such approaches combined



with substrate walking methods<sup>189</sup> have enabled the utilisation of halogenases in the selective decoration of increasingly more complex substrates, as well as stabilising enzymes to operate at higher temperatures.<sup>190</sup> Engineering of the NHFeHals has also enabled their application to a broader series of substrates (see Section 2.2.4).<sup>178,179</sup> Though the evolution of a small number of tryptophan halogenases for utilisation for halogenation of a broader series of substrates has proved useful, other approaches have considered the natural diversity of the halogenases. The broad range of halogenated structures that have been noted, over 5000 to date,<sup>13</sup> together with the variety of organisms from which they are produced, indicates that many diverse halogenases remain to be discovered. *In silico* discovery approaches have been utilised to mine for various halogenases within uncured genomic deposits, leading to the discovery of more substrate flexible halogenases from diverse sources including viruses.<sup>47,170</sup> Discovery of novel halogenases remains at the forefront of biocatalytic halogenation research. For effective biocatalysis, good amounts of soluble protein are required. The Lewis group have effectively improved heterologous production of RebF through including a cleavable maltose binding protein (MBP) tag, to enhance solubility of the enzyme, and co-expressing both the halogenase and the concomitant reductase, RebF with the chaperonins GroEL/ES, in addition to developing a cofactor recycling system.<sup>126</sup> Advances on FADH<sub>2</sub> regeneration have recently been achieved by replacing NADH with synthetic mimics.<sup>191</sup> A photocatalytic approach has been successfully applied to the VHPOs. The Gulder group cleverly demonstrated that by illuminating their biocatalytic system with 455 nm light, and including sacrificial electron donors, cofactor recycling could be addressed, and a ready supply of reduced cofactor afforded in a relatively inexpensive manner.<sup>192,193</sup> An approach that has made a significant difference to the utilisation of FDHs is the application of cross-linked enzyme aggregates (CLEA). FDHs generally have poor stability, impacting their catalytic longevity and utility in scaled up reactions. The robust nature and solvent tolerance of Cross-Linked Enzyme Crystals was first reported in the early 1990s, but the technique required the enzymes of interest to first be crystallised.<sup>194</sup> The application of CLEA was first demonstrated for PenAcyase, an aminoacylase utilised in the generation of ampicillin, where CLEA were generated from precipitating, rather than crystallising, the enzyme prior to cross-linking.<sup>195</sup> By using a CombiCLEAS approach of 1 g of L-tryptophan to 1.8 g of L-7-Br-tryptophan could be achieved with complete conversion within an 8 days period.<sup>130</sup> To achieve this, the lysate from *Escherichia coli* (*E. coli*) (6 L), that had been grown with a RebH expression vector, was combined with PrnF, the flavin reductase from *Pseudomonas fluorescens*, and an alcohol dehydrogenase from *Rhodococcus* sp., that had been precipitated with ammonium sulfate, was then cross-linked with 0.5%/wt glutaraldehyde to form the CombiCLEAS.

#### 4. Site selective C–H activation through hyphenation of enzymatic halogenation and synthetic diversification

The opportunity to combine enzymatic halogenation with chemical modification, in particular catalytic cross-coupling, is exciting and provides a much-needed tool for molecule

constructions. This concept, first reported in 2010, with the tryptophan halogenase PrnA being utilised to give access to a new to nature haloindolic natural product, which was subsequently functionalised through Suzuki–Miyaura cross-coupling,<sup>196</sup> has been gaining momentum. This section comprehensively reviews the current state-of-the-art and reflects on technologies enabling the hyphenation.

##### 4.1 Aqueous compatible cross-coupling methodologies.

Palladium-catalysed cross-coupling reactions play a central role in organic chemistry. Their utility in the formation of C–C and carbon–heteroatom bonds under mild conditions with a wide range of organic halide and nucleophilic reagents, has led to them being one of the most utilised series of reactions in the pharmaceutical industry.<sup>197</sup> As with almost all enzymes, the halogenases require an aqueous environment for effective operation. To effectively couple enzymatic halogenation with synthetic diversification, aqueous compatible reaction conditions are generally required. Palladium-catalysed cross-coupling reactions have largely utilized traditional organic solvents. Standard conditions for these coupling methodologies typically use a mixed organic/aqueous base solvent system, hence the development of water solubilising ligands have been highly enabling. The first example of a water-soluble palladium catalyst for cross-coupling reactions being reported by Casalnuovo,<sup>198</sup> and significant effort has been devoted to the development of hydrophilic palladium/ligand complexes for aqueous-phase catalysis since this seminal work, as well as active hydrophobic catalysts and ligand-free species based on palladium nanoparticles.<sup>199–203</sup> The development of mild cross coupling conditions for the modulation of challenging and or functionally sensitive compounds is advancing, enabling the diversification and tuning of small molecules, natural products and biomolecules including nucleic acids and peptides. To this end, aqueous compatible and water soluble ligands such as TPPTS **274**,<sup>199</sup> and the sterically demanding and electron-rich TXPTS **275**,<sup>199</sup> <sup>s</sup>SPhos **276**,<sup>204</sup> <sup>s</sup>XPhos **277**<sup>204</sup> and disodium 2-aminopyrimidine-4,6-diol **278**<sup>205–207</sup> have proved valuable (Fig. 10). Additionally, nitrogen based ligands enabling aqueous compatible cross-couplings such as tetramethylguanidine **279** and 2-(dimethylamino)-pyrimidine-4,6-diol **280**, both reported to have low cell toxicity, are beginning

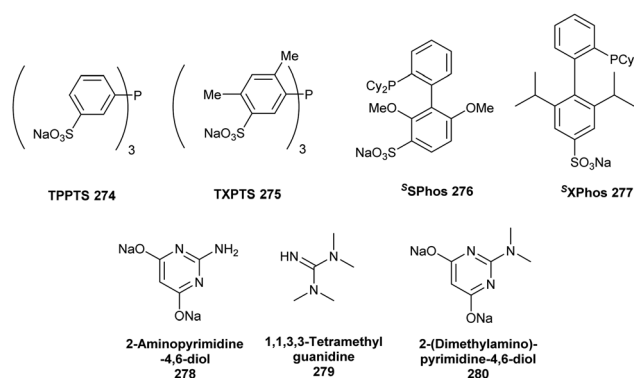


Fig. 10 Selected water-soluble ligands used for aqueous Suzuki–Miyaura cross-coupling reactions in the presence of halogenase enzymes or even living systems.



to facilitate tandem halogenation and cross-coupling strategies in living systems.<sup>208</sup>

Specifically, with the goal of enabling C–H activation through enzymatic halogenation and chemical cross-coupling, our team has developed series of aqueous cross-coupling procedures, enabling the cross-coupling to be compatible with molecules decorated with sensitive functional groups, as well as with enzymes and even in live cell environments. We have initiated our development of these catalytic conditions by focussing on challenging free and unprotected halotryptophans. Our choice has been directed by three factors:

- The availability of a well-characterised series of tryptophan halogenases that may be utilised in standalone format in conjunction with a biosynthetic pathway to access halotryptophans and synchronously new-to-nature halotryptophan containing natural products;
- The utility of tryptophan as a fluorescence tag and a moiety that governs protein folding, and the attraction to being able to tune these properties;
- And the challenge that free tryptophan confers, rendering it a suitably stimulating test bed. We reasoned that by developing conditions that were suitable to address the cross-coupling of free halotryptophan, a small molecule with the propensity to coordinate to the catalyst and render it inactive, such conditions could be readily applied to other more tractable moieties. To this end, we developed a range of aqueous conditions enabling the Suzuki–Miyaura,<sup>209</sup> Heck,<sup>210</sup> Sonogashira,<sup>211</sup> keto-arylation<sup>212</sup> and Buchwald–Hartwig<sup>213</sup> diversification of free unprotected tryptophan, peptides, and natural products (Scheme 13).

The development of mild aqueous chemistries has facilitated chemo-enzymatic and GenoChemetic approaches. Site selective C–H activation is enabled by a halogenase and synthetic diversification carried out, in some cases synchronously, through aqueous compatible Pd mediated catalysis. More reactive aryl iodides and bromides can be coupled under much milder conditions than their less activated chloro counterparts.

**4.2 Chemo-enzymatic approaches.** Hyphenation of enzymatic halogenation and synthetic derivatisation has been enabled by steering, where possible, to enzymatic bromination (or even iodination) over enzymatic chlorination, employing combinations of stabilised or partitioned enzymes, and use of mild and aqueous cross-coupling methodologies. There is burgeoning interest in employing such strategies as powerful methods for late-stage diversification on ever more complex molecules. The approach is also utilised to enable a fast fluorescence assay for enzyme screening and for directed evolution. A few enzymes, such as RadH have been shown capable of halogenating 7-hydroxycoumarin. An enhancement in fluorescence occurs when 7-hydroxycoumarin ( $\lambda_{\text{max}}$  325 nm) is converted to 8-chloro-7-hydroxycoumarin ( $\lambda_{\text{ex}}$  386 nm/ $\lambda_{\text{em}}$  456 nm).<sup>144</sup> In most halogenation reactions, a clear change in fluorescence is not always observable, and cross-coupling chemistry has been employed to provide a fluorescence readout that may be used to assay halogenase activity. Sewald reported the utilisation of tryptophan 5-, 6- and 7-halogenases PyrH, ThaI, and RebH for bromination of L-tryptophan

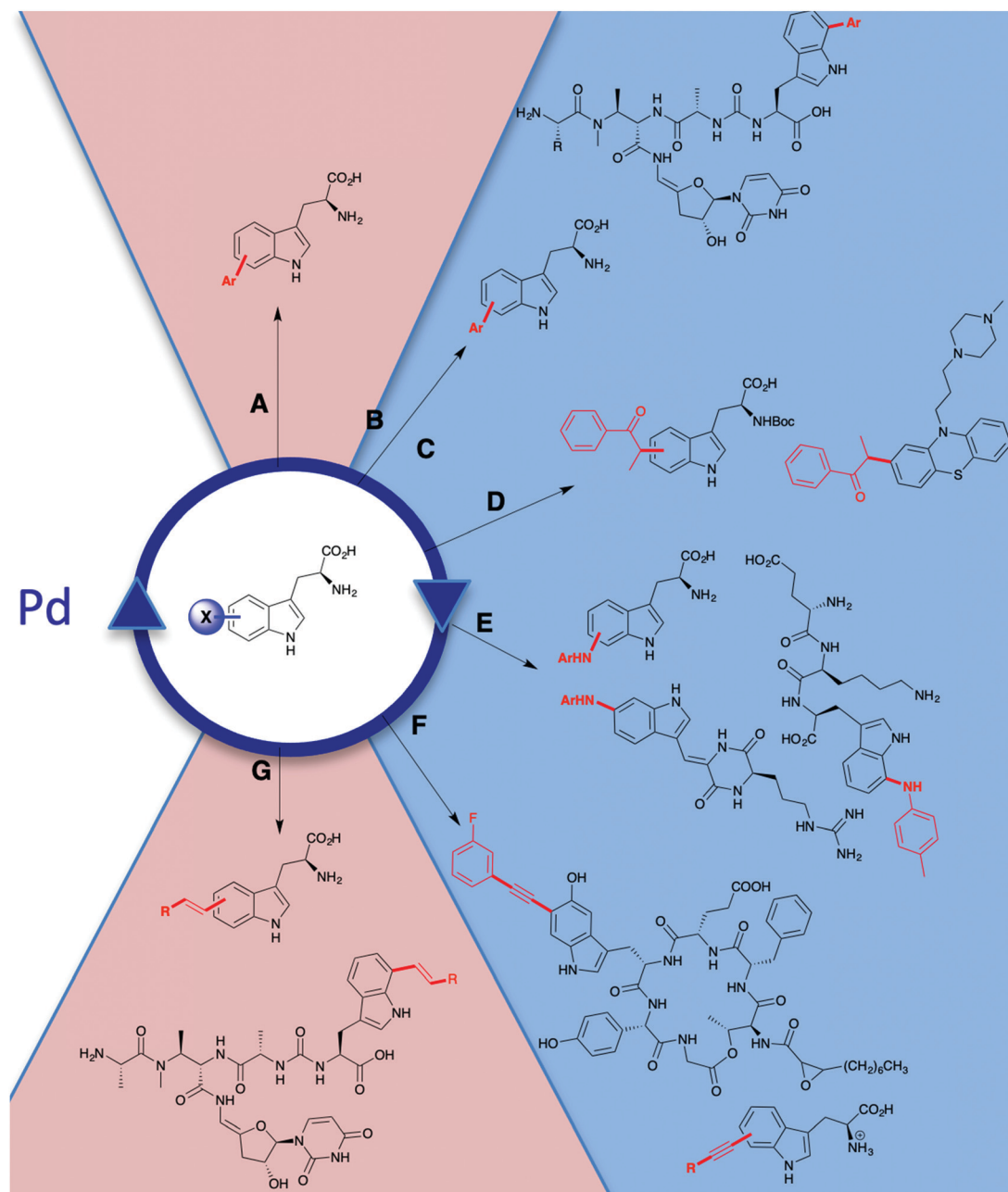
substrate.<sup>214,215</sup> The team then employed aqueous cross-coupling conditions. To achieve cross-coupling of the bromo-tryptophan product within the crude lysate, the Pd-catalyst loading was further increased to 50 mol%. By reaction with the boronic acid of aniline, the products (3'-aminophenyl)-tryptophans were reported to have a  $\lambda_{\text{ex}}$  300 nm/ $\lambda_{\text{em}}$  430 nm (Scheme 14B). This technology was utilised for microtitre plate screening enabling rapid analysis of ThaI mutants that had been generated by error-prone PCR.<sup>215</sup> The Sewald team also cleverly employed CLEA technology (see Section 3) to enable bromination of tryptophan by these same halogenases PyrH, ThaI, and RebH. Not only did the utilisation of enzyme aggregates afford stabilisation, but it also provided an opportunity for readily filtering off the enzymes ahead of Suzuki–Miyaura diversification of the resultant halotryptophans. By utilising a three-step approach, organic solvents required for solubilisation of several of the substrates could be used, with Boc protection being carried out in a final step, thereby easing purification.<sup>214</sup> In a similar manner, the Sewald group developed a three-step one pot reaction utilising RebH and Mizoroki–Heck chemistry for the further diversification of tryptophans. Again, by using their combi-CLEA system in a stepwise manner (Scheme 14D), the harsher reaction conditions required for Heck diversification could be utilised and enabled the generation of C-7 substituted styryl-tryptophan with a  $\lambda_{\text{em}}$  of 485 nm ( $\lambda_{\text{ex}}$  360 nm), and Stokes shift of 7159 cm<sup>−1</sup>.<sup>216</sup>

This larger Stokes' shift renders this technology potentially very useful for molecule tagging and enzyme assay, as well as valuable for molecule diversification. Even when stabilised as CLEA, enzymes and Pd catalysts can have poor compatibility. To achieve a synchronous one-pot enzymatic halogenation and chemical cross-coupling, and to avoid the additional filtration step, Micklefield and co-workers elegantly partitioned CLEA stabilised PyrH, RebH, RadH and SttH from the Pd catalyst through use of a polydimethylsiloxane (PDMS) membrane. Whilst the brominated enzymatic product could pass through the PMDS, it was reported to be impervious to the enzymes and cofactors. The cross-coupled product, once formed, remained in the chemo-catalytic chamber (Scheme 14G).<sup>218</sup>

Though the majority of chemo-enzymatic approaches published so far have focussed on tryptophans and other indolic systems, this is predominantly a reflection of the most extensively studied FDHs. Through substrate walking methodology, Jared Lewis and co-workers have impressively evolved RebH to work on substrates that are chemically distinct from its natural tryptophan substrate. Utilisation of evolved halogenases enabled bromination of a diverse series of substrates including tryptolines, sterically bulky substituted anilines as well as carvediol and pindolol, two clinically utilised beta-blockers. Conditions to enable Suzuki–Miyaura, Buchwald–Hartwig amination and alkoxylation were developed and applied mostly to the chlorinated and brominated intermediates as a component of a crude extract, resulting in high yields of the respective products **284–289** (Scheme 13D–F).<sup>189</sup> With the discovery of more substrate diverse halogenases, engineering of known halogenases, enzyme stabilisation and partitioning strategies and the development of milder reaction conditions,



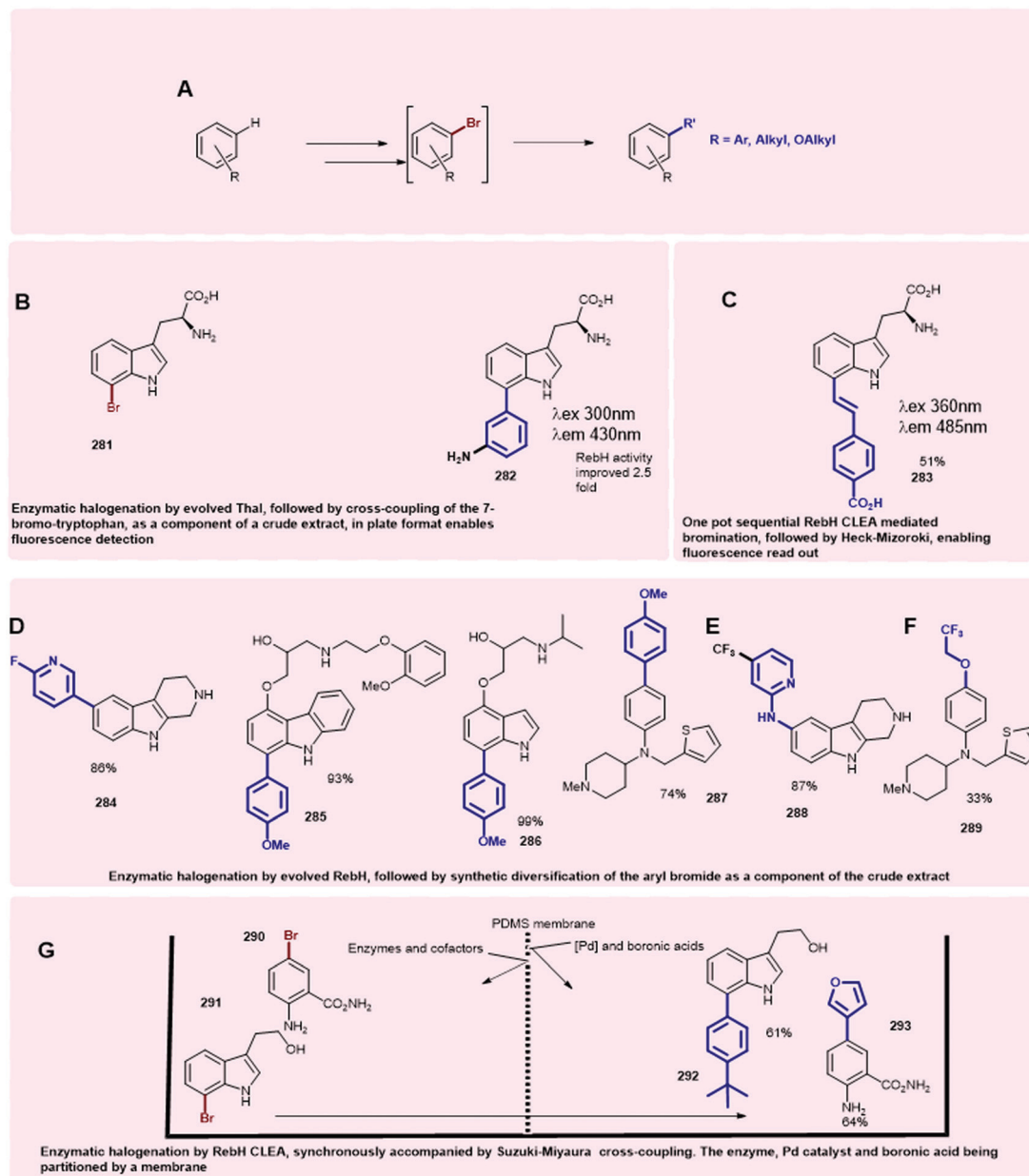




**Scheme 13** Mild aqueous conditions enabling the cross-coupling of free unprotected halotryptophan and halotryptophans embedded into natural products, biomolecules and bioactive compounds. Colour code: Blue background: conditions enable processing of aryl-chlorides in addition to more reactive aryl-bromides, Red background: least activated aryl halide processed is Ar-Br under these conditions (broad substrate scope and extension of chemistry to other systems not shown here). (A) Suzuki-Miyaura cross-coupling on bromotryptophans: phenylboronic acid (1.1 equiv.),  $\text{Na}_2\text{PdCl}_4$  (2.5 mol%), TXPTS (25 mol%),  $\text{K}_2\text{CO}_3$  (5 equiv.), water, 40–80 °C, 70–90% isolated yields.<sup>209</sup> (B) Suzuki-Miyaura cross-coupling on chlorotryptophans, and 7-chlorotryptophan pacidamycin: phenylboronic acid (1.5 equiv.),  $\text{Na}_2\text{PdCl}_4$  (5 mol%), SPhos (10 mol%),  $\text{K}_2\text{CO}_3$  (5 equiv.), water:acetonitrile 80 °C, on purified material and crude extract 60–70% isolated yields.<sup>217</sup> (C) Suzuki-Miyaura cross-coupling of bromotryptophans, *p*-tolylboronic acid (3 equiv.),  $\text{Pd}(\text{OAc})_2$ -[2-amino-pyrimidine-4,6-diyl]<sub>2</sub> (5 mol%),  $\text{K}_2\text{CO}_3$  (6 equiv.), 45 °C, 54–85% isolated yields and 7-bromotryptophan pacidamycin in living cultures: reaction carried out in optimised culture medium at 37 °C.<sup>206</sup> (D) Ketoarylation of *N*-Boc-bromotryptophans, and small water-soluble chloro and bromo pharmacophores: ketone (4 equiv.),  $[\text{Pd}(\text{DtBPF})\text{Cl}_2]$  (DtBPF = 1,1'-bis(di-*tert*-butylphosphino)ferrocene) (2 mol%), NaOH (4 equiv.), dioxane : water 1 : 1, 60 °C, 45–94% isolated yields.<sup>212</sup> (E) Buchwald-Hartwig diversification of chloro and bromotryptophans, peptides containing bromotryptophan and the bromotryptophan containing natural product baretin: aniline (2 equiv.),  $[\text{Pd}^t\text{Bu-XPhosG1}]$  (5 mol%), KOH (4 equiv.), dioxane : water 1 : 1, 100 °C, 8–48 min, 31–75% isolated yields.<sup>213</sup> (F) Sonogashira diversification of bromotryptophans, peptides containing bromotryptophan and the new to nature bromotryptophan containing natural product cystargamide: alkyne (1 equiv.),  $[\text{PdCl}_2(\text{CH}_3\text{CN})_2]$  (5 mol%), sXPhos (15 mol%),  $\text{Cs}_2\text{CO}_3$  (2.5 equiv.), acetonitrile : water 1 : 1, 100 °C, 25–97% isolated yields.<sup>211</sup> (G) Heck diversification of bromotryptophans, peptides containing bromotryptophan 7-bromotryptophan containing pacidamycin: alkene (1.5 equiv.),  $\text{Na}_2\text{PdCl}_4$  (10 mol%), TXPTS (23 mol%),  $\text{Na}_2\text{CO}_3$  (4 equiv.), acetonitrile : water 1 : 3, 90 °C, 51–95% isolated yields.<sup>210</sup>





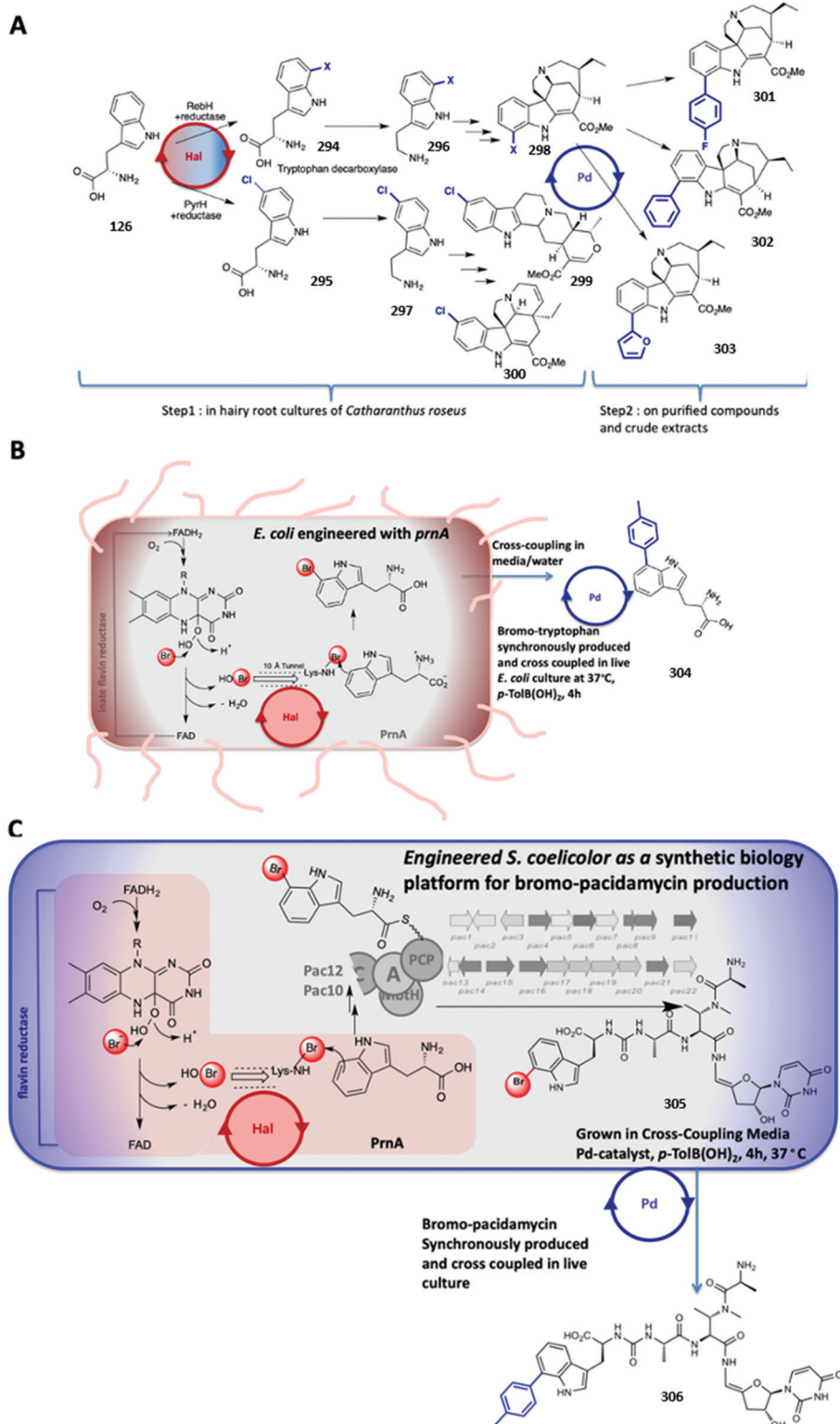


**Scheme 14** Hyphenation of halogenation and cross-coupling to achieve C–H activation, (A) general concept in which the halogenase operates first, and then the cross-coupling is carried out on the halogenated material as a component of the crude extract. This is carried out stepwise in (B–F), enabling access to a broad series of compounds, and synchronously in G where the halogenation and cross-coupling events are spatially separated through application of a membrane. (B) Bromination of L-tryptophan by Thal, followed by Suzuki–Miyaura cross-coupling of the brominated product as a component of the crude lysate.<sup>215</sup> The product of cross-coupling ( $\lambda_{\text{ex}}$  300 nm/ $\lambda_{\text{em}}$  430 nm) is used in screening the directed evolution of RebH, resulting in a 2.5-fold increase in enzymatic activity. Cross-coupling conditions: 3-aminophenylboronic acid (10 equiv.),  $\text{Na}_2\text{PdCl}_4$  (50 mol%),  $^5\text{Sphos}$  (150 mol%),  $\text{K}_3\text{PO}_4$  (15 equiv.), 95 °C.<sup>215</sup> (C) One-pot bromination of L-tryptophan by RebH-CLEA, followed by Mizoroki–Heck cross-coupling, in a stepwise fashion to afford C-7 substituted 7-(4-carboxystyryl)tryptophan with a  $\lambda_{\text{em}}$  of 485 nm ( $\lambda_{\text{ex}}$  360 nm). Cross-coupling conditions: 4-carboxystyrene (5 equiv.),  $\text{PdOAc}_2$  (0.1–0.2 equiv.), TPPTS (0.3–0.6 equiv.),  $\text{K}_2\text{CO}_3$  (5 equiv.), water, degassed under argon, 100 °C.<sup>189</sup> (D–F) Sequential enzymatic bromination by evolved RebH variants, followed by (D) Suzuki–Miyaura, (E) Buchwald–Hartwig, (F) alkoxylation of tryptophan as a component of a crude extract. Conditions D:  $\text{ArB}(\text{OH})_2$  (1.5 equiv.),  $\text{PdOAc}_2$  (0.05 equiv.),  $^5\text{Sphos}$  (0.05 equiv.), iPrOH: phosphate buffer (170 mM, pH 8.5) 1:1, 90 °C. (E)  $\text{ArNH}_2$  (3 equiv.),  $\text{PdOAc}_2$  (0.03 equiv.), BrettPhos (0.03 equiv.),  $\text{NaOt-Bu}$  (6 equiv.), dioxane, 100 °C. (F)  $\text{CF}_3\text{CH}_2\text{OH}$  (2 equiv.),  $[(\text{allyl})\text{PdCl}]_2$  (0.005 equiv.), RockPhos (0.015 equiv.),  $\text{Cs}_2\text{CO}_3$  (2 equiv.), toluene, 90 °C.<sup>189</sup> (G) One-pot, synchronous halogenation and Suzuki–Miyaura cross-coupling by PDMS membrane separated RebH-CLEA and Pd catalyst. Cross-coupling conditions:  $\text{PdOAc}_2$  (10 mol%), 2-(dimethylamino)-pyrimidine-4,6-diol (20 mol%), aryl boronic acid (5 equiv.) and CsF (10 equiv.), r.t. overnight then 80 °C.<sup>211</sup>

the portfolio of functionalisable systems is set to expand significantly.

FDHs are not the only halogenating enzymes to be merged with palladium catalysed cross-coupling. V-HPOs and HPOs





**Scheme 15** GenoChemetic approaches to C–H functionalisation of small metabolites and complex natural products. Here, a halogenase is installed to complement an existing metabolic/biosynthetic pathway, the halogenated natural product is diversified through application of cross-coupling chemistries. (A) Two step approach: hairy root cultures of *Catharanthus roseus* are engineered to include RebH, PyrH and their concomitant reductases. A series of chlorinated and brominated analogues are produced, and cross coupling has been carried out on purified compounds and on the metabolites as components of crude extracts, enabling access to a wide portfolio of functionalised molecules.<sup>224</sup> Cross-coupling conditions: arylboronic acid (3 equiv.), Pd(OAc)<sub>2</sub> (5 mol%), Sphos (13 mol%), K<sub>3</sub>PO<sub>4</sub>, *n*-butanol, 100 °C, 10–60 min. (B and C) One pot approach, pathways constructed in heterologous hosts, with halogenation (mediated by PrnA)<sup>225</sup> and cross-coupling occurring synchronously in the presence of the living *E. coli* and *Streptomyces coelicolor*,<sup>222</sup> respectively.



have been successfully used in various chemoenzymatic halocyclization reactions,<sup>219,220</sup> and recently this has been merged with homogenous metal catalysis for further diversification. Deska and coworkers enabled the one-pot combination of allenol halocyclization using CPO from *Caldariomyces fumago* and glucose oxidase from *Aspergillus niger* (GOx) with SPhos catalysed Suzuki–Miyaura and Sonogashira coupling to diversify a range of dihydrofurans. Excitingly, a nanobiohybrid was also generated using GOx as a support for palladium nanoparticles, which led to good yields of stepwise halocyclization and Suzuki coupling.<sup>221</sup>

**4.3 GenoChemetic approaches.** Whilst in an enzymatic or whole cell biotransformation where the enzymes or cells are exogenously supplemented with substrates, in a GenoChemetic approach, the substrate and tagged intermediates are fully biosynthesised by an engineered organism. GenoChemetic processes provide an expeditious route to generating series of analogues of natural products. Semi-synthesis, though a useful tool, is limited by the innate functional groups naturally present within the molecule. GenoChemetics addresses this. It involves the introduction of a gene to complement a biosynthetic pathway, such that a chemically reactive and orthogonal handle is introduced into the natural product. This handle enables synthetic diversification of the natural product. There are so far three reports of the utilisation of halogenases within GenoChemetic systems. In each of these systems, tryptophan halogenases have been employed, with RebH, PyrH, and PrnA being installed into either pathway engineered or wild-type strains enabling the production of chlorinated or brominated analogues of the complex natural product scaffolds.

The halogenated metabolites have been subjected to cross-coupling chemistries. In the first example, the tryptophan halogenase PrnA was installed, through use of an integrative vector, into the genome of *Streptomyces coeruleorubidis*. It was demonstrated that there was no need to introduce PrnA's concomitant flavin reductase, as the organism's own reductases sufficed. A chlorinated analogue of pacidamycin, a uridyl peptide antibiotic, was produced at a titre of  $\sim 1 \text{ mg L}^{-1}$ . In spite of bearing series of sensitive functional groups, this metabolite could be readily cross-coupled using the moderately mild Suzuki–Miyaura conditions that were developed. The cross-coupling was carried out both on the purified material and on material as a component of a crude extract (Scheme 15). The latter was preferable, as the more lipophilic product that resulted was much more readily amenable to purification.<sup>217</sup> The generation of brominated analogues would, of course, facilitate cross-coupling under even milder conditions. However, *Streptomyces coeruleorubidis* was found to be sensitive to bromide salts, and the brominated analogue cannot readily be accessed through use of this strain. *Streptomyces coelicolor* showed good tolerance to bromide salts in a screen of series of potential heterologous expression hosts, and so the pacidamycin biosynthetic gene cluster, identified previously, was ported into this strain conferring it with the ability to produce pacidamycins. The gene encoding the halogenase PrnA was engineered into the strain and ability to produce bromo-pacidamycin was confirmed.

In order to enable bromometabolite production and cross-coupling synchronously, mild cross-coupling reactions that enabled derivatisation of 7-Br-pacidamycin D **305** in the presence of the living culture were carried out. In a similar manner, *E. coli* was engineered with the ability to make bromotryptophan and its cross-coupling to **306** affected in the living fermentative culture.<sup>222</sup> A similar system was designed in hairy root cultures of *Catharanthus roseus* in which they were engineered to include RebH, PyrH and their concomitant reductases. This led to the production of a range of chlorinated and brominated analogues of the plant's monoterpene indole alkaloid natural products **299–301**, which were successfully extracted and coupled in non-aqueous Suzuki–Miyaura couplings, either as crude extracts or purified compounds (Scheme 15A).<sup>223,224</sup>

## Conclusions

Though enzyme mediated halogenation was initially considered to be nothing more than an artefact or a rare event, today over 30 enzymes capable of mediating such chemistry have been structurally and biochemically characterised (Table 2), and an even greater number subject to *in vitro* and *in vivo* investigation. Halogenated compounds play a major role in the construction of molecules *via* cross-coupling methodologies, as well as making up a large percentage of medicinally and agrochemically relevant molecules. In a world that seeks more selective, greener and more sustainable solutions to molecule manufacture, halogenases hold an opportunity. They also offer potential for the vectorisation of fragments for SAR. Whilst the kinetic analyses of the halogenases is rather potted, drawing this data together is useful in identifying potential halogenases for development and use in biotransformations. Whilst the HPOs are capable of mediating halogenation of large and sterically demanding substrates, as most neither bind the substrate nor the electrophilic species, such systems generally lack regiocontrol. A few intriguing departures from this trend have been identified from marine systems, indicating that there is much more to learn about these enzymes, their operation and evolution. New variant A FDHs with reasonable catalytic efficiency and broad substrate scopes have been revealed in the last two years, enabling regioselective chlorination, bromination and even iodination on sterically and electronically demanding aromatic systems, even prior to rational redesign or evolution. Excitingly, in recent years the range of variant A NHFeHals that enable the controlled halogenation, including stereoselective halogenation, have been identified; clearly, there are many more such systems awaiting discovery. Inspired by such systems, it may be possible to generate artificial architectures that, perhaps *via* haloamines, position and deliver electrophilic halogenating agents to substrates held in the required position to affect desired regiochemistries.

Halogenation introduces a chemically reactive, chemically orthogonal moiety into a compound. This single atom modification, that can be genetically/enzymatically installed has the potential to act as a superior tag to azide or alkylne



modifications. In parallel with halogenase discovery, a variety of aqueous cross-coupling reactions compatible with sensitive functionalities in natural products and biomolecules have been developed. Some of these reactions are even biorthogonal. The opportunities for building pathways and circuits of SynBio-SynChem present themselves.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We thank ERC GenoChemetics (FP7/2007-2013/ERC consolidator grant GCGXC grant agreement no. 614779 RJMG) for funding, BBSRC Follow on Funding, Synthetic Biology Highlight award, RSE Enterprise Fellowship (SVS), China Scholarship Council and EPSRC CRITICAT EP/L016419/1 for studentship support (Y. Z. & S. M. respectively).

## References

- 1 P. Jeschke, *Pest Manage. Sci.*, 2017, **73**, 1053–1066.
- 2 P. Jeschke, R. Nauen, M. Schindler and A. Elbert, *J. Agric. Food Chem.*, 2011, **59**, 2897–2908.
- 3 Z. Xu, Z. Yang, Y. Liu, Y. Lu, K. Chen and W. Zhu, *J. Chem. Inf. Model.*, 2014, **54**, 69–78.
- 4 M. C. Ford and P. S. Ho, *J. Med. Chem.*, 2016, **59**, 1655–1670.
- 5 C. W. Lindsley, *ACS Chem. Neurosci.*, 2017, **8**, 1635–1636.
- 6 B. E. Blass, *Basic Principles of Drug Discovery and Development*, 2015, pp. 499–529.
- 7 R. Whitfield and F. Brown, *The Benefits of Chlorine Chemistry in Pharmaceuticals in the United States and Canada*, 2016.
- 8 E. Parisini, P. Metrangolo, T. Pilati, G. Resnati and G. Terraneo, *Chem. Soc. Rev.*, 2011, **40**, 2267–2278.
- 9 J. Ruiz-Cabello, B. P. Barnett, P. A. Bottomley and J. W. M. Bulte, *NMR Biomed.*, 2011, **24**, 114–129.
- 10 J. M. Widmark, *Bayl. Univ. Med. Cent. Proc.*, 2007, **20**, 408–417.
- 11 J. G. Speight, *Environmental Organic Chemistry for Engineers*, 2017, pp. 87–151.
- 12 D. A. Petrone, J. Ye and M. Lautens, *Chem. Rev.*, 2016, **116**, 8003–8104.
- 13 G. W. Gribble, *Environ. Chem.*, 2015, **12**, 396–405.
- 14 P. D. Shaw and L. P. Hager, *J. Am. Chem. Soc.*, 1959, **81**, 6527–6528.
- 15 J. R. Beckwith, R. Clark and L. P. Hager, *J. Biol. Chem.*, 1963, **9**, 3086–3090.
- 16 J. Delaney, E. Clarke, D. Hughes and M. Rice, *Drug Discovery Today*, 2006, **11**, 839–845.
- 17 P. Maienfisch and R. G. Hall, *Chimia*, 2004, **58**, 93–99.
- 18 R. J. Milne and K. L. Goa, *Drugs*, 1991, **41**, 450–477.
- 19 M. Baumann and I. R. Baxendale, *Beilstein J. Org. Chem.*, 2013, **9**, 2265–2319.
- 20 P. Oakley, *J. Pharm. Res. Int.*, 2018, **21**, 1–9.
- 21 A. Timmins and S. P. de Visser, *Catalysts*, 2018, **8**, 314.
- 22 J. Reedijk and E. Bouwman, *Bioinorganic Catalysis*, CRC Press, Boca Raton, 2nd edn, 1999.
- 23 H. B. Dunford, A. M. Lambeir, M. A. Kashem and M. Pickard, *Arch. Biochem. Biophys.*, 1987, **252**, 292–302.
- 24 L. C. Blasiak and C. L. Drennan, *Acc. Chem. Res.*, 2009, **42**, 147–155.
- 25 M. Fritzsche, V. Vergopoulos and D. Rehder, *Inorg. Chim. Acta*, 1993, **211**, 11–16.
- 26 E. de Boer and R. Wever, *J. Biol. Chem.*, 1988, **263**, 12326–12332.
- 27 G. J. Colpas, B. J. Hamstra, J. W. Kampf and V. L. Pecoraro, *J. Am. Chem. Soc.*, 1996, **118**, 3469–3478.
- 28 C. C. McLauchlan, H. A. Murakami, C. A. Wallace and D. C. Crans, *J. Inorg. Biochem.*, 2018, **186**, 267–279.
- 29 R. Morris and L. P. Hager, *J. Biol. Chem.*, 1966, **8**, 1763–1768.
- 30 P. Bernhardt, T. Okino, J. M. Winter, A. Miyanaga and B. S. Moore, *J. Am. Chem. Soc.*, 2011, **133**, 4268–4270.
- 31 S. Diethelm, R. Teufel, L. Kaysser and B. S. Moore, *Angew. Chem., Int. Ed.*, 2014, **53**, 11023–11026.
- 32 B. S. Moore, *Synlett*, 2018, 401–409.
- 33 T. Nakano, K. Aisaka, R. Katsumata, M. Hasegawa and T. Dai, *Biosci., Biotechnol., Biochem.*, 1995, **59**, 1099–1106.
- 34 S. Kirner, P. E. Hammer, D. S. Hill, A. Altmann, I. Fischer, L. J. Weislo, M. Lanahan and J. M. Ligon, *J. Bacteriol.*, 1998, **180**, 1939–1943.
- 35 S. Keller, T. Wage, K. Hohaus, M. Hölzer, E. Eichhorn and P. Kh, *Angew. Chem., Int. Ed.*, 2000, **39**, 2300–2302.
- 36 S. Buedenbender, S. Rachid, R. Müller and G. E. Schulz, *J. Mol. Biol.*, 2009, **385**, 520–530.
- 37 C. Dong, S. Flecks, S. Unversucht, C. Haupt, K. H. Van Pée and J. H. Naismith, *Science*, 2005, **309**, 2216–2219.
- 38 E. Yeh, L. C. Blasiak, A. Koglin, C. L. Drennan and C. T. Walsh, *Biochemistry*, 2007, **46**, 1284–1292.
- 39 S. Flecks, E. P. Patallo, X. Zhu, A. J. Ernyei, G. Seifert, A. Schneider, C. Dong, J. H. Naismith and K. H. Van Pée, *Angew. Chem., Int. Ed.*, 2008, **47**, 9533–9536.
- 40 E. Yeh, L. J. Cole, E. W. Barr, J. M. Bollinger, D. P. Ballou and C. T. Walsh, *Biochemistry*, 2006, **45**, 7904–7912.
- 41 M. C. Andorfer and J. C. Lewis, *Annu. Rev. Biochem.*, 2018, **87**, 159–185.
- 42 M. C. Andorfer, H. J. Park, J. Vergara-Coll and J. C. Lewis, *Chem. Sci.*, 2016, **7**, 3720–3729.
- 43 E. Yeh, L. C. Blasiak, A. Koglin, C. L. Drennan and C. T. Walsh, *Biochemistry*, 2007, **46**, 1284–1292.
- 44 K. Podzelinska, R. Latimer, A. Bhattacharya, L. C. Vining, D. L. Zechel and Z. Jia, *J. Mol. Biol.*, 2010, **397**, 316–331.
- 45 V. Agarwal, A. A. El Gamal, K. Yamanaka, D. Poth, R. D. Kersten, M. Schorn, E. E. Allen and B. S. Moore, *Nat. Chem. Biol.*, 2014, **10**, 640–647.
- 46 A. E. Fraley, M. Garcia-Borràs, A. Tripathi, D. Khare, E. V. Mercado-Marin, H. Tran, Q. Dan, G. P. Webb,





- K. R. Watts, P. Crews, R. Sarpong, R. M. Williams, J. L. Smith, K. N. Houk and D. H. Sherman, *J. Am. Chem. Soc.*, 2017, **139**, 12060–12068.
- 47 D. S. Gkotsi, H. Ludewig, S. V. Sharma, J. A. Connolly, J. Dhaliwal, Y. Wang, W. P. Unsworth, R. J. K. Taylor, M. M. W. McLachlan, S. Shanahan, J. H. Naismith and R. J. M. Goss, *Nat. Chem.*, 2019, **11**, 1091–1097.
- 48 A. El Gamal, V. Agarwal, S. Diethelm, I. Rahman, M. A. Schorn, J. M. Sneed, G. V. Louie, K. E. Whalen, T. J. Mincer, J. P. Noel, V. J. Paul and B. S. Moore, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 3797–3802.
- 49 O. Dym and D. Eisenberg, *Protein Sci.*, 2001, **10**, 1712–1728.
- 50 J. Ahlert, E. Shepard, N. Lomovskaya, E. Zazopoulos, A. Staffa, B. O. Bachmann, K. Huang, L. Fonstein, A. Csisny, R. E. Whitwam, C. M. Farnet and J. S. Thorson, *Science*, 2002, **297**, 1173–1176.
- 51 S. Mori, A. H. Pang, N. T. Chandrika, S. Garneau-tsodikova and O. V. Tsodikov, *Nat. Commun.*, 2019, 1–11.
- 52 S. S. Gao, N. Naowarojna, R. Cheng, X. Liu and P. Liu, *Nat. Prod. Rep.*, 2018, **35**, 792–837.
- 53 W. S. Aik, M. A. McDonough, A. Thalhammer, R. Chowdhury and C. J. Schofield, *Curr. Opin. Struct. Biol.*, 2012, **22**, 691–700.
- 54 M. L. Matthews, C. M. Krest, E. W. Barr, F. H. Vaillancourt, C. T. Walsh, M. T. Green, C. Krebs and J. M. Bollinger, *Biochemistry*, 2009, **48**, 4331–4343.
- 55 D. P. Galonić, E. W. Barr, C. T. Walsh, J. M. Bollinger and C. Krebs, *Nat. Chem. Biol.*, 2007, **3**, 113–116.
- 56 M. L. Matthews, C. S. Neumann, L. A. Miles, T. L. Grove, S. J. Booker, C. Krebs, C. T. Walsh and J. M. Bollinger, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 17723–17728.
- 57 S. M. Pratter, K. M. Light, E. I. Solomon and G. D. Straganz, *J. Am. Chem. Soc.*, 2014, **136**, 9385–9395.
- 58 F. H. Vaillancourt, J. Yin and C. T. Walsh, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 10111–10116.
- 59 X. Liu, *In Vitro Analysis of Cyanobacterial Nonheme Iron-Dependent Aliphatic Halogenases WelO5 and AmbO5*, Elsevier Inc., 1st edn, 2018, vol. 604.
- 60 D. P. Galonić, F. H. Vaillancourt and C. T. Walsh, *J. Am. Chem. Soc.*, 2006, **128**, 3900–3901.
- 61 L. Gu, B. Wang, A. Kulkarni, T. W. Geders, R. V. Grindberg, L. Gerwick, K. Hkansson, P. Wipf, J. L. Smith, W. H. Gerwick and D. H. Sherman, *Nature*, 2009, **459**, 731–735.
- 62 M. L. Hillwig and X. Liu, *Nat. Chem. Biol.*, 2014, **10**, 921–923.
- 63 D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. G. Hamilton and C. D. Murphy, *Nature*, 2002, **416**, 279.
- 64 A. S. Eustáquio, F. Pojer, J. P. Noel and B. S. Moore, *Nat. Chem. Biol.*, 2008, **4**, 69–74.
- 65 D. O'Hagan and H. Deng, *Chem. Rev.*, 2015, **115**, 634–649.
- 66 H. M. Senn, D. O'Hagan and W. Thiel, *J. Am. Chem. Soc.*, 2005, **127**, 13643–13655.
- 67 C. Dong, F. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan and J. H. Naismith, *Nature*, 2004, **427**, 561–565.
- 68 H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson and J. B. Spencer, *Angew. Chem., Int. Ed.*, 2006, **45**, 759–762.
- 69 S. Thompson, M. Onega, S. Ashworth, I. N. Fleming, J. Passchier and D. O'Hagan, *Chem. Commun.*, 2015, **51**, 13542–13545.
- 70 M. C. Walker, B. W. Thuronyi, L. K. Charkoudian, B. Lowry, C. Khosla and M. C. Y. Chang, *Science*, 2013, **341**, 1089–1094.
- 71 D. Milbredt, E. P. Patallo and K. H. Van Pée, *ChemBioChem*, 2014, **15**, 1011–1020.
- 72 J. Zeng and J. Zhan, *Biotechnol. Lett.*, 2011, **33**, 1607–1613.
- 73 J. Lee, J. Kim, H. Kim, E. J. Kim, H. J. Jeong, K. Y. Choi and B. G. Kim, *ChemBioChem*, 2020, **21**, 1446–1452.
- 74 K. Lingkon and J. J. Bellizzi, *ChemBioChem*, 2020, **21**, 1121–1128.
- 75 Z. Guo, P. Li, G. Chen, C. Li, Z. Cao, Y. Zhang, J. Ren, H. Xiang, S. Lin, J. Ju and Y. Chen, *J. Am. Chem. Soc.*, 2018, **140**, 18009–18015.
- 76 S. Zehner, A. Kotzsch, B. Bister, R. D. Süssmuth, C. Méndez, J. A. Salas and K. H. Van Pée, *Chem. Biol.*, 2005, **12**, 445–452.
- 77 J. Domergue, D. Erdmann, A. Fossey-Jouenne, J.-L. Petit, A. Debar, V. de Berardinis, C. Vergne-Vaxelaire and A. Zaparucha, *AMB Express*, 2019, **9**, 175.
- 78 A. Lang, S. Polnick, T. Nicke, P. William, E. P. Patallo, J. H. Naismith and K. H. Van Pée, *Angew. Chem., Int. Ed.*, 2011, **50**, 2951–2953.
- 79 S. A. Shepherd, C. Karthikeyan, J. Latham, A. W. Struck, M. L. Thompson, B. R. K. Menon, M. Q. Styles, C. Levy, D. Leys and J. Micklefield, *Chem. Sci.*, 2015, **6**, 3454–3460.
- 80 S. A. Shepherd, B. R. K. Menon, H. Fisk, A. W. Struck, C. Levy, D. Leys and J. Micklefield, *ChemBioChem*, 2016, **17**, 821–824.
- 81 A. C. Moritzer, H. Minges, T. Prior, M. Frese, N. Sewald and H. H. Niemann, *J. Biol. Chem.*, 2019, **294**, 2529–2542.
- 82 J. R. Heemstra and C. T. Walsh, *J. Am. Chem. Soc.*, 2008, **130**, 14024–14025.
- 83 M. L. Hillwig, Q. Zhu, K. Ittiamornkul and X. Liu, *Angew. Chem., Int. Ed.*, 2016, **55**, 5780–5784.
- 84 M. E. Neugebauer, K. H. Sumida, J. G. Pelton, J. L. McMurphy, J. A. Marchand and M. C. Y. Chang, *Nat. Chem. Biol.*, 2019, **15**, 1009–1016.
- 85 C. Zhao, S. Yan, Q. Li, H. Zhu, Z. Zhong, Y. Ye, Z. Deng and Y. Zhang, *Angew. Chem., Int. Ed.*, 2020, **59**, 9478–9484.
- 86 R. D. Libby, N. S. Rotberg, T. Emerson, T. C. White, G. M. Yen, S. H. Friedman, N. S. Sun and R. Goldowski, *J. Biol. Chem.*, 1989, **264**, 15284–15292.
- 87 X. Zhu, D. A. Robinson, A. R. McEwan, D. O'Hagan and J. H. Naismith, *J. Am. Chem. Soc.*, 2007, **129**, 14597–14604.
- 88 E. Bitto, Y. Huang, C. A. Bingman, S. Singh, J. S. Thorson and G. N. Phillips, *Proteins: Struct., Funct., Bioinf.*, 2007, **70**, 289–293.
- 89 C. B. Poor, M. C. Andorfer and J. C. Lewis, *ChemBioChem*, 2014, **15**, 1286–1289.
- 90 X. Zhu, W. De Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K. H. van Pée and J. H. Naismith, *J. Mol. Biol.*, 2009, **391**, 74–85.
- 91 A. H. Pang, S. Garneau-Tsodikova and O. V. Tsodikov, *J. Struct. Biol.*, 2015, **192**, 349–357.





- 92 B. R. K. Menon, J. Latham, M. S. Dunstan, E. Brandenburger, U. Klemstein, D. Leys, C. Karthikeyan, M. F. Greaney, S. A. Shepherd and J. Micklefield, *Org. Biomol. Chem.*, 2016, **14**, 9354–9361.
- 93 P. R. Neubauer, C. Widmann, D. Wibberg, L. Schröder, M. Frese, T. Kottke, J. Kalinowski, H. H. Niemann and N. Sewald, *PLoS One*, 2018, **13**, 1–21.
- 94 H. Luhavaya, R. Sigrist, J. R. Chekan, S. M. K. McKinnie and B. S. Moore, *Angew. Chem., Int. Ed.*, 2019, **58**, 8394–8399.
- 95 A. C. Moritzer and H. H. Niemann, *Protein Sci.*, 2019, **28**, 2112–2118.
- 96 A. Phintha, K. Prakinee, A. Jaruwat, N. Lawan, S. Visitsatthawong, C. Kantiwiriyanitch, W. Songsunthong, D. Trisvirat, P. Chenprakhon, A. Mulholland, K.-H. van Pée, P. Chitnumsub and P. Chaiyen, *J. Biol. Chem.*, 2021, **296**, 100068.
- 97 A. Messerschmidt and R. Wever, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 392–396.
- 98 S. Macedo-Ribeiro, W. Hemrika, R. Renirie, R. Wever and A. Messerschmidt, *J. Biol. Inorg. Chem.*, 1999, **4**, 209–219.
- 99 A. Messerschmidt, L. Prade and R. Wever, *Biol. Chem.*, 1997, **378**, 309–316.
- 100 S. De Macedo-Ribeiro, R. Renirie, R. Wever and A. Messerschmidt, *Biochemistry*, 2008, **47**, 929–934.
- 101 M. Weyand, H. J. Hecht, M. Kieß, M. F. Liaud, H. Vilter and D. Schomburg, *J. Mol. Biol.*, 1999, **293**, 595–611.
- 102 J. N. Carter-Franklin, J. D. Parrish, R. A. Tschirret-Guth, R. D. Little and A. Butler, *J. Am. Chem. Soc.*, 2003, **125**, 3688–3689.
- 103 D. Wischang, M. Radlow and J. Hartung, *Dalton Trans.*, 2013, **42**, 11926–11940.
- 104 M. N. Isupov, A. R. Dalby, A. A. Brindley, Y. Izumi, T. Tanabe, G. N. Murshudov and J. A. Littlechild, *J. Mol. Biol.*, 2000, **299**, 1035–1049.
- 105 E. Garcia-Rodriguez, T. Ohshiro, T. Aibara, Y. Izumi and J. Littlechild, *J. Biol. Inorg. Chem.*, 2005, **10**, 275–282.
- 106 J. B. Fournier, E. Rebuffet, L. Delage, R. Grijol, L. Meslet-Cladière, J. Rzonca, P. Potin, G. Michel, M. Czjzek and C. Leblanc, *Appl. Environ. Microbiol.*, 2014, **80**, 7561–7573.
- 107 A. Frank, C. J. Seel, M. Groll and T. Gulder, *ChemBioChem*, 2016, **17**, 2028–2032.
- 108 M. Sundaramoorthy, J. Turner and T. L. Poulos, *Structure*, 1995, **3**, 1367–1377.
- 109 K. Ku, W. Blankenfeldt, J. Turner and I. Schlichting, *J. Biol. Chem.*, 2006, **281**, 23990–23998.
- 110 C. Wong, D. G. Fujimori, C. T. Walsh and C. L. Drennan, *J. Am. Chem. Soc.*, 2009, **131**, 4872–4879.
- 111 L. C. Blasiak, F. H. Vaillancourt, C. T. Walsh and C. L. Drennan, *Nature*, 2006, **440**, 368–371.
- 112 S. L. Cobb, H. Deng, A. R. McEwan, J. H. Naismith, D. O'Hagan and D. A. Robinson, *Org. Biomol. Chem.*, 2006, **4**, 1458–1460.
- 113 S. Thompson, S. A. McMahon, J. H. Naismith and D. O'Hagan, *Bioorg. Chem.*, 2016, **64**, 37–41.
- 114 H. Deng, L. Ma, N. Bandaranayaka, Z. Qin, G. Mann, K. Kyremeh, Y. Yu, T. Shepherd, J. H. Naismith and D. O'Hagan, *ChemBioChem*, 2014, **15**, 364–368.
- 115 S. Marumo and H. Hattori, *Planta*, 1970, **90**, 208–211.
- 116 M. Ishibashi, Y. Li, M. Sato and J. Kobayashi, *Nat. Prod. Lett.*, 1994, **4**, 293–296.
- 117 S. Rachid, D. Krug, B. Kunze, I. Kochems, M. Scharfe, T. M. Zabriskie, H. Blöcker and R. Müller, *Chem. Biol.*, 2006, **13**, 667–681.
- 118 W. Jiang, J. R. Heemstra, R. R. Forseth, C. S. Neumann, S. Manaviazar, F. C. Schroeder, K. J. Hale and C. T. Walsh, *Biochemistry*, 2011, **50**, 6063–6072.
- 119 D. R. M. Smith, A. R. Uria, E. J. N. Helfrich, D. Milbredt, K. H. Van Pée, J. Piel and R. J. M. Goss, *ACS Chem. Biol.*, 2017, **12**, 1281–1287.
- 120 K. Hohaus, A. Altmann, W. Burd, I. Fischer, P. E. Hammer, D. S. Hill, J. M. Ligon and K. Van Pée, *Angew. Chem., Int. Ed.*, 2012, **36**, 2012–2013.
- 121 D. G. Fujimori, S. Hrvatin, C. S. Neumann, M. Strieker, M. A. Marahiel and C. T. Walsh, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16498–16503.
- 122 D. E. Williams, J. Davies, B. O. Patrick, H. Botttriell, T. Tarling, M. Roberge and R. J. Andersen, *Org. Lett.*, 2008, **10**, 3501–3504.
- 123 F. Y. Chang and S. F. Brady, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 2478–2483.
- 124 C. Sánchez, L. Zhu, A. F. Braña, A. P. Salas, J. Rohr, C. Méndez and J. A. Salas, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 461–466.
- 125 Z. Liu, L. Ma, L. Zhang, W. Zhang, Y. Zhu, Y. Chen, W. Zhangb and C. Zhang, *Org. Biomol. Chem.*, 2019, **17**, 1053–1057.
- 126 J. T. Payne, M. C. Andorfer and J. C. Lewis, *Angew. Chem., Int. Ed.*, 2013, **52**, 5271–5274.
- 127 J. T. Payne, C. B. Poor and J. C. Lewis, *Angew. Chem., Int. Ed.*, 2015, **54**, 4226–4230.
- 128 F. Y. Chang and S. F. Brady, *ChemBioChem*, 2014, **15**, 815–821.
- 129 C. Carrasco, M. Facompré, J. D. Chisholm, D. L. Van Vranken, W. D. Wilson and C. Bailly, *Nucleic Acids Res.*, 2002, **30**, 1774–1781.
- 130 M. Frese and N. Sewald, *Angew. Chem., Int. Ed.*, 2015, **54**, 298–301.
- 131 X. M. Chu, C. Wang, W. Liu, L. L. Liang, K. K. Gong, C. Y. Zhao and K. L. Sun, *Eur. J. Med. Chem.*, 2019, **161**, 101–117.
- 132 F. Prati, C. Bergamini, R. Fato, O. Soukup, J. Korabecny, V. Andrisano, M. Bartolini and M. L. Bolognesi, *ChemMedChem*, 2016, 1284–1295.
- 133 G. A. Castillon, N. R. Adames, C. H. Rosello, H. S. Seidel, M. S. Longtine, J. A. Cooper and R. A. Heil-Chapdelaine, *Curr. Biol.*, 2003, **13**, 654–658.
- 134 O. Bilyk, E. Brötz, B. Tokovenko, A. Bechthold, T. Paululat and A. Luzhetskyy, *ACS Chem. Biol.*, 2016, **11**, 241–250.
- 135 N. A. Magarvey, Z. Q. Beck, T. Golakoti, Y. Ding, U. Huber, T. K. Hemscheidt, D. Abelson, R. E. Moore and D. H. Sherman, *ACS Chem. Biol.*, 2006, **1**, 766–779.
- 136 T. Kittila, C. Kittel, J. Tailhades, D. Butz, M. Schoppet, A. Buttner, R. J. A. Goode, R. B. Schittenhelm, K.-H. van Pee,



- R. D. Sussmuth, W. Wohlleben, M. J. Cryle and E. Stegmann, *Chem. Sci.*, 2017, **8**, 5992–6004.
- 137 S. Lin, S. G. Van Lanen and B. Shen, *J. Am. Chem. Soc.*, 2007, **129**, 12432–12438.
- 138 S. Rachid, M. Scharfe, H. Blöcker, K. J. Weissman and R. Müller, *Chem. Biol.*, 2009, **16**, 70–81.
- 139 X. Y. Jia, Z. H. Tian, L. Shao, X. D. Qu, Q. F. Zhao, J. Tang, G. L. Tang and W. Liu, *Chem. Biol.*, 2006, **13**, 575–585.
- 140 S. Cadel-Six, C. Dauga, A. M. Castets, R. Rippka, C. Bouchier, N. Tandeau de Marsac and M. Welker, *Mol. Biol. Evol.*, 2008, **25**, 2031–2041.
- 141 S. Stinchi, L. Carrano, A. Lazzarini, M. Feroggio, A. Grigoletto, M. Sosio and S. Donadio, *FEMS Microbiol. Lett.*, 2006, **256**, 229–235.
- 142 Y. Ge, G. Wang, J. Jin, T. Liu, X. Ma, Z. Zhang, Y. Ge, G. Wang, J. Jin, T. Liu, X. Ma, Z. Zhang and T. Geng, *J. Org. Chem.*, 2020, **85**, 8673–8682.
- 143 J. Zeng, A. K. Lytle, D. Gage, S. J. Johnson and J. Zhan, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 1001–1003.
- 144 B. R. K. Menon, E. Brandenburger, H. H. Sharif, U. Klemstein, S. A. Shepherd, M. F. Greaney and J. Micklefield, *Angew. Chem., Int. Ed.*, 2017, **56**, 11841–11845.
- 145 Y. H. Chooi, R. Cacho and Y. Tang, *Chem. Biol.*, 2010, **17**, 483–494.
- 146 J. M. Winter, M. Sato, S. Sugimoto, G. Chiou, N. K. Garg, Y. Tang and K. Watanabe, *J. Am. Chem. Soc.*, 2012, **134**, 17900–17903.
- 147 T. Zhu, X. Cheng, Y. Liu, Z. Deng and D. You, *Metab. Eng.*, 2013, **19**, 69–78.
- 148 M. T. Nielsen, J. B. Nielsen, D. C. Anyagwu, D. K. Holm, K. F. Nielsen, T. O. Larsen and U. H. Mortensen, *PLoS One*, 2013, **8**, e72871.
- 149 P. Chankhamjon, D. Boettger-Schmidt, K. Scherlach, B. Urbansky, G. Lackner, D. Kalb, H. M. Dahse, D. Hoffmeister and C. Hertweck, *Angew. Chem., Int. Ed.*, 2014, **53**, 13409–13413.
- 150 L. Kong, Q. Wang, Z. Deng and D. You, *Appl. Environ. Microbiol.*, 2020, **86**, 1–14.
- 151 I. Wynands and K. H. Van Pée, *FEMS Microbiol. Lett.*, 2004, **237**, 363–367.
- 152 S. M. K. McKinnie, Z. D. Miles and B. S. Moore, *Methods in Enzymology*, Elsevier Inc., 1st edn, 2018, vol. 604, pp. 405–424.
- 153 S. M. K. McKinnie, Z. D. Miles, P. A. Jordan, T. Awakawa, H. P. Pepper, L. A. M. Murray, J. H. George and B. S. Moore, *J. Am. Chem. Soc.*, 2018, **140**, 17840–17845.
- 154 M. C. Andorfer, J. E. Grob, C. E. Hajdin, J. R. Chael, P. Siuti, J. Lilly, K. L. Tan and J. C. Lewis, *ACS Catal.*, 2017, **7**, 1897–1904.
- 155 Z. D. Miles, S. Diethelm, H. P. Pepper, D. M. Huang, J. H. George and B. S. Moore, *Nat. Chem.*, 2017, **9**, 1235–1242.
- 156 C. Wannstedt, D. Rotella and J. F. Siuda, *Bull. Environ. Contam. Toxicol.*, 1990, 282–287.
- 157 M. C. R. Franssen and H. C. van der Plas, *Bioorg. Chem.*, 1987, **15**, 59–70.
- 158 P. Yaipakdee and L. W. Robertson, *Phytochemistry*, 2001, **57**, 341–347.
- 159 D. Wischang and J. Hartung, *Tetrahedron*, 2012, **68**, 9456–9463.
- 160 F. Xu, A. Merkley, D. Yu and J. Zhan, *Tetrahedron Lett.*, 2016, **57**, 5262–5265.
- 161 Q. Yan, B. Philmus, J. H. Chang and J. E. Loper, *eLife*, 2017, **6**, 1–24.
- 162 S. M. Mantovani and B. S. Moore, *J. Am. Chem. Soc.*, 2013, **135**, 18032–18035.
- 163 I. Höfer, M. Crüsemann, M. Radzom, B. Geers, D. Flachshaar, X. Cai, A. Zeeck and J. Piel, *Chem. Biol.*, 2011, **18**, 381–391.
- 164 P. C. Dorrestein, E. Yeh, S. Garneau-Tsodikova, N. L. Kelleher and C. T. Walsh, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 13843–13848.
- 165 X. Zhang and R. J. Parry, *Antimicrob. Agents Chemother.*, 2007, **51**, 946–957.
- 166 K. Yamanaka, K. S. Ryan, T. A. M. Gulder, C. C. Hughes and B. S. Moore, *J. Am. Chem. Soc.*, 2012, **134**, 12434–12437.
- 167 N. Itoh, A. K. M. Q. Hasan, Y. Izumi and H. Yamada, *Eur. J. Biochem.*, 1988, **172**, 477–484.
- 168 J. Hartung, Y. Dumont, M. Greb, D. Hach, F. Köhler, H. Schulz, M. Casný, D. Rehder and H. Vilter, *Pure Appl. Chem.*, 2009, **81**, 1251–1264.
- 169 R. Vázquez-Duhalt, M. Ayala and F. J. Márquez-Rocha, *Phytochemistry*, 2001, **58**, 929–933.
- 170 B. F. Fisher, H. M. Snodgrass, K. A. Jones, M. C. Andorfer and J. C. Lewis, *ACS Cent. Sci.*, 2019, **5**, 1844–1856.
- 171 M. Ueki, D. P. Galonić, F. H. Vaillancourt, S. Garneau-Tsodikova, E. Yeh, D. A. Vosburg, F. C. Schroeder, H. Osada and C. T. Walsh, *Chem. Biol.*, 2006, **13**, 1183–1191.
- 172 S. M. Pratter, J. Ivkovic, R. Birner-Gruenberger, R. Breinbauer, K. Zangger and G. D. Straganz, *ChemBioChem*, 2014, **15**, 567–574.
- 173 N. Sitachitta, J. Rossi, M. A. Roberts, W. H. Gerwick, M. D. Fletcher and C. L. Willis, *J. Am. Chem. Soc.*, 1998, **120**, 7131–7132.
- 174 Z. Chang, N. Sitachitta, J. V. Rossi, M. A. Roberts, P. M. Flatt, J. Jia, D. H. Sherman and W. H. Gerwick, *J. Nat. Prod.*, 2004, **67**, 1356–1367.
- 175 F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. E. O'Connor and C. T. Walsh, *Nature*, 2005, **436**, 1191–1194.
- 176 Q. Zhu, M. L. Hillwig, Y. Doi and X. Liu, *ChemBioChem*, 2016, **17**, 466–470.
- 177 M. L. Hillwig, Q. Zhu and X. Liu, *ACS Chem. Biol.*, 2014, **9**, 372–377.
- 178 S. Duewel, L. Schmermund, T. Faber, K. Harms, V. Srinivasan, E. Meggers and S. Hoebenreich, *ACS Catal.*, 2019, **10**, 1272–1277.
- 179 T. Hayashi, M. Ligibel, E. Sager, M. Voss, J. Hunziker, K. Schroer, R. Snajdrova and R. Buller, *Angew. Chem., Int. Ed.*, 2019, 18535–18539.
- 180 Q. Zhu and X. Liu, *Beilstein J. Org. Chem.*, 2017, **13**, 1168–1173.



- 181 C. Y. Kim, A. J. Mitchell, C. M. Glinkerman, F. S. Li, T. Pluskal and J. K. Weng, *Nat. Commun.*, 2020, **11**, 1–7.
- 182 A. J. Mitchell, N. P. Dunham, J. A. Bergman, B. Wang, Q. Zhu, W. C. Chang, X. Liu and A. K. Boal, *Biochemistry*, 2017, **56**, 441–444.
- 183 J. A. Marchand, M. E. Neugebauer, M. C. Ing, C. I. Lin, J. G. Pelton and M. C. Y. Chang, *Nature*, 2019, **567**, 420–424.
- 184 J. N. Carter-Franklin and A. Butler, *J. Am. Chem. Soc.*, 2004, **126**, 15060–15066.
- 185 M. Pirae, R. L. White and L. C. Vining, *Microbiology*, 2004, **150**, 85–94.
- 186 P. Chankhamjon, Y. Tsunematsu, M. Ishida-Ito, Y. Sasa, F. Meyer, D. Boettger-Schmidt, B. Urbansky, K. D. Menzel, K. Scherlach, K. Watanabe and C. Hertweck, *Angew. Chem., Int. Ed.*, 2016, **55**, 11955–11959.
- 187 M. C. Andorfer and J. C. Lewis, *Annu. Rev. Biochem.*, 2018, **87**, 159–185.
- 188 H. Ludewig, S. Molyneux, S. Ferrinho, K. Guo, R. Lynch, D. S. Gkotsi and R. J. M. Goss, *Curr. Opin. Struct. Biol.*, 2020, **65**, 51–60.
- 189 L. J. Durak, J. T. Payne and J. C. Lewis, *ACS Catal.*, 2016, **6**, 1451–1454.
- 190 H. Minges, C. Schnepel, D. Böttcher, M. S. Weiß, J. Sproß, U. T. Bornscheuer and N. Sewald, *ChemCatChem*, 2020, **12**, 818–831.
- 191 M. Ismail, L. Schroeder, M. Frese, T. Kottke, F. Hollmann, C. E. Paul and N. Sewald, *ACS Catal.*, 2019, **9**, 1389–1395.
- 192 C. J. Seel, A. Králík, M. Hacker, A. Frank, B. König and T. Gulder, *ChemCatChem*, 2018, **10**, 3960–3963.
- 193 C. J. Seel and T. Gulder, *ChemBioChem*, 2019, **20**, 1871–1897.
- 194 N. L. S. Clair and M. A. Navia, *J. Am. Chem. Soc.*, 1992, **114**, 7314–7316.
- 195 L. Cao, F. Van Rantwijk and R. A. Sheldon, *Org. Lett.*, 2000, **2**, 1361–1364.
- 196 A. D. Roy, S. Grischow, N. Cairns and R. J. M. Goss, *J. Am. Chem. Soc.*, 2010, **132**, 12243–12245.
- 197 C. C. C. J. Seechurn, M. O. Kitching, T. J. Colacot and V. Snieckus, *Angew. Chem., Int. Ed.*, 2012, 5062–5085.
- 198 A. L. Casalnuovo and J. C. Calabrese, *J. Am. Chem. Soc.*, 1990, **112**, 4324–4330.
- 199 E. C. Western, J. R. Daft and K. H. Shaughnessy, *J. Org. Chem.*, 2003, **68**, 6767–6774.
- 200 T. E. Schmid, D. C. Jones, O. Songis, O. Diebolt, M. R. L. Furst, A. M. Z. Slawin and C. S. J. Cazin, *Dalton Trans.*, 2013, **42**, 7345–7353.
- 201 G. Hervé, G. Sartori, G. Enderlin, G. MacKenzie and C. Len, *RSC Adv.*, 2014, **4**, 18558–18594.
- 202 B. H. Lipshutz, T. B. Petersen and A. R. Abela, *Org. Lett.*, 2008, **10**, 1333–1336.
- 203 A. Fihri, M. Bouhrara, B. Nekoueishahraki, J. M. Basset and V. Polshettiwar, *Chem. Soc. Rev.*, 2011, **40**, 5181–5203.
- 204 K. W. Anderson and S. L. Buchwald, *Angew. Chem., Int. Ed.*, 2005, **44**, 6173–6177.
- 205 J. M. Chalker, C. S. C. Wood and B. G. Davis, *J. Am. Chem. Soc.*, 2009, 16346–16347.
- 206 T. Willemse, K. Van Imp, R. J. M. Goss, H. W. T. Van Vlijmen, W. Schepens, B. U. W. Maes and S. Ballet, *ChemCatChem*, 2015, 2055–2070.
- 207 J. H. Li, X. D. Zhang and Y. X. Xie, *Eur. J. Org. Chem.*, 2005, 4256–4259.
- 208 Z. Gao, V. Gouverneur and B. G. Davis, *J. Am. Chem. Soc.*, 2013, **135**, 13612–13615.
- 209 A. D. Roy, R. J. M. Goss, G. K. Wagner and M. Winn, *Chem. Commun.*, 2008, 4831–4833.
- 210 R. J. M. Goss, C. Pubill-Ulldemolins, S. Sharma, C. Cartmell, J. Zhao and P. Cárdenas, *Chem. – Eur. J.*, 2019, **25**, 10866–10875.
- 211 M. J. Corr, S. V. Sharma, C. Pubill-Ulldemolins, R. T. Bown, P. Poirot, D. R. M. Smith, C. Cartmell, A. Abou Fayad and R. J. M. Goss, *Chem. Sci.*, 2017, **8**, 2039–2046.
- 212 E. Marelli, Y. Renault, S. V. Sharma, S. P. Nolan and R. J. M. Goss, *Chem. – Eur. J.*, 2017, **23**, 3832–3836.
- 213 Y. J. G. Renault, R. Lynch, E. Marelli, S. V. Sharma, C. Pubill-Ulldemolins, J. A. Sharp, C. Cartmell, P. Cárdenas and R. J. M. Goss, *Chem. Commun.*, 2019, **55**, 13653–13656.
- 214 M. Frese, C. Schnepel, H. Minges, H. Voß, R. Feiner and N. Sewald, *ChemCatChem*, 2016, **8**, 1799–1803.
- 215 C. Schnepel, H. Minges, M. Frese and N. Sewald, *Angew. Chem., Int. Ed.*, 2016, **55**, 14159–14163.
- 216 H. Gruß, C. Belu, L. M. Bernhard, A. Merschel and N. Sewald, *Chem. – Eur. J.*, 2019, **25**, 5880–5883.
- 217 A. D. Roy, S. Grischow, N. Cairns and R. J. M. Goss, *J. Am. Chem. Soc.*, 2010, **132**, 12243–12245.
- 218 M. F. Greaney, J. Micklefield, B. R. K. Menon, S. A. Shepherd, H. H. Sharif, J. Latham and J.-M. Henry, *Nat. Commun.*, 2016, **7**, 1–8.
- 219 S. H. H. Younes, F. Tieves, D. Lan, Y. Wang, P. Süß, H. Brundiek, R. Wever and F. Hollmann, *ChemSusChem*, 2019, 97–101.
- 220 J. Naapuri, J. D. Rolfes, J. Keil, C. Manzuna Sapu and J. Deska, *Green Chem.*, 2017, **19**, 447–452.
- 221 J. M. Naapuri, G. A. Åberg, J. M. Palomo and J. Deska, *ChemCatChem*, 2021, **13**, 763–769.
- 222 S. V. Sharma, X. Tong, C. Pubill-Ulldemolins, C. Cartmell, E. J. A. Bogosyan, E. J. Rackham, E. Marelli, R. B. Hamed and R. J. M. Goss, *Nat. Commun.*, 2017, **8**, 1–9.
- 223 W. Runguphan, X. Qu and S. E. O'Connor, *Nature*, 2010, **468**, 461–467.
- 224 W. Runguphan and S. E. O'Connor, *Org. Lett.*, 2013, **15**, 2850–2853.
- 225 E. J. Rackham, S. Grischow, A. E. Ragab, S. Dickens and R. J. M. Goss, *ChemBioChem*, 2010, **11**, 1700–1709.
- 226 Y. Xiao, S. Li, S. Niu, L. Ma, G. Zhang, H. Zhang, G. Zhang, J. Ju and C. Zhang, *J. Am. Chem. Soc.*, 2011, **133**, 1092–1105.

